

# Blind map level systematics cleaning: a quadratic estimator approach

Joel Williams,<sup>a</sup> Nialh McCallum,<sup>a</sup> Aditya Rotti,<sup>a</sup>  
Daniel B. Thomas,<sup>a,b</sup> Richard Battye<sup>a</sup> and Michael L. Brown<sup>a</sup>

<sup>a</sup>Jodrell Bank Centre for Astrophysics, Department of Physics and Astronomy, School of Natural Sciences, University of Manchester, Manchester, M13 9PL, U.K.

<sup>b</sup>School of Physics and Astronomy, Queen Mary University of London, London, E1 4NS, U.K.

E-mail: [christopher.williams-4@manchester.ac.uk](mailto:christopher.williams-4@manchester.ac.uk),  
[nialh.mccallum@manchester.ac.uk](mailto:nialh.mccallum@manchester.ac.uk), [aditya.rotti@manchester.ac.uk](mailto:aditya.rotti@manchester.ac.uk),  
[daniel.thomas-2@manchester.ac.uk](mailto:daniel.thomas-2@manchester.ac.uk), [richard.battye@manchester.ac.uk](mailto:richard.battye@manchester.ac.uk),  
[m.l.brown@manchester.ac.uk](mailto:m.l.brown@manchester.ac.uk)

Received January 25, 2021

Accepted June 9, 2021

Published July 9, 2021

**Abstract.** We present the first detailed case study using quadratic estimators (QE) to diagnose and remove systematics present in observed Cosmic Microwave Background (CMB) maps. In this work we focus on the temperature to polarization leakage. We use an iterative QE analysis to remove systematics, in analogy to de-lensing, recovering the primordial B-mode signal and the systematic maps. We introduce a new Gaussian filtering scheme crucial to stable convergence of the iterative cleaning procedure and validate with comparisons to semi-analytical forecasts. We study the limitations of this method by examining its performance both on idealized simulations and on more realistic, non-ideal simulations, where we assume varying de-lensing efficiencies. Finally, we quantify the systematic cleaning efficiency by presenting a likelihood analysis on the tensor to scalar ratio,  $r$ , and demonstrate that the blind cleaning results in an un-biased measurement of  $r$ , reducing the systematic induced B-mode power by nearly two orders of magnitude.

**Keywords:** CMBR detectors, CMBR experiments, CMBR polarisation, gravitational waves and CMBR polarization

**ArXiv ePrint:** [2101.09097](https://arxiv.org/abs/2101.09097)



---

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Distortions of the CMB</b>	<b>3</b>
2.1	Additional distortion terms	4
<b>3</b>	<b>Quadratic estimators</b>	<b>5</b>
3.1	Quadratic estimator for the spin-2 $\gamma$ field	5
3.2	The reconstruction noise	8
3.2.1	Impact of de-lensing on reconstruction noise	9
3.3	Efficient real space estimators	10
<b>4</b>	<b>Iterative cleaning</b>	<b>10</b>
4.1	Optimally de-contaminating the CMB maps	11
4.1.1	Gaussian filters	12
4.2	Forecasting the reconstruction noise and $C_l^{BB}$ floors	12
<b>5</b>	<b>Simulating temperature to polarization leakage</b>	<b>13</b>
5.1	Systematic-differential gain	13
5.2	Differential gain $r$ -bias	14
5.3	Scan strategy	15
5.4	Simulation	17
5.5	Realistic systematic spectra	18
<b>6</b>	<b>Reconstructing and removing the temperature to polarization leakage</b>	<b>19</b>
6.1	The cosmic variance limits of the blind cleaning algorithm	19
6.2	Employing blind systematic cleaning for a non-ideal experiment	23
<b>7</b>	<b>Recovering the tensor-to-scalar ratio</b>	<b>25</b>
<b>8</b>	<b>Conclusions</b>	<b>26</b>
<b>A</b>	<b>Geometric identity</b>	<b>28</b>

---

## 1 Introduction

The Cosmic Microwave Background (CMB) intensity and polarization are key observables for cosmology. The frontiers of cosmology have been pushed back by progressively more sensitive measurement of the CMB delivered by Cosmic Background Explorer (*COBE*) [1], the Wilkinson Microwave Anisotropy Probe (*WMAP*) [2], and *Planck* [3]. Future surveys, including satellite and ground-based, will measure the CMB sky with unprecedented sensitivities, with the primordial B-mode polarization spectrum as one of the primary targets. With these ever increasing sensitivities, precision control over systematics and their removal will become increasingly important.

Robust measurement of the primordial B-mode signals will require overcoming a number of analysis challenges, including foreground removal, removal of the weak-lensing B-mode

signal and potential contamination from instrument systematic effects. Recent studies have shown that the upcoming experiments in principle have sufficient sensitivity and frequency coverage for robust recovery of B-mode signal corresponding to  $r \sim 10^{-3}$  [e.g. 4, 5]. Similarly other studies have examined how well B-mode skies may be de-lensed [e.g. 6]. In this work we focus our attention on systematics from instruments, specifically examining them using quadratic estimators (QE).

Common approaches to the removal of instrument systematics require complex modeling and prior knowledge of the instrument itself. An appealing aspect of a QE approach is that it in principle allows a largely agnostic approach to dealing with instrument systematics. That is, the effect of systematics on the CMB data can be modeled as a set distortions to the CMB data. Reconstructing these distortions using a QE does not require prior knowledge or modeling of the instrument sourcing the distortions. QE cleaning and reconstruction is therefore a promising complimentary approach to traditional systematic modeling techniques.

Previous works have suggested the use of QE as a method to quantify the level of systematics in CMB maps [e.g 7], formulating QE for a variety of instrumental systematics. These included gain fluctuations, differential systematics, and instrumental polarisation rotation, to list a few among a much longer list of possible instrument systematics. QEs are most commonly employed in reconstruction of the lensing potential map [e.g. 8] but have also been used to constrain cosmological birefringence [e.g. 9–11]. QE studies frequently draw from understanding gained from lensing reconstruction. We improve upon previous studies of QEs as applied to systematic effects by considering an estimator in the full sky regime, accounting for realistic scan patterns, and testing whether certain aspects of the conventional wisdom from lensing studies apply. We consider two scenarios throughout this work: (i) a no lensing, noise, and beam free scenario which we refer to as the “ideal case”, and (ii) a scenario we refer to as the “more realistic, non-ideal case” that includes the effect of lensing on the CMB spectra, a gaussian white noise of  $w_{TT}^{-1} = 2.7 \mu\text{K arcmin}$  and a full width at half maximum of  $\theta_{\text{FWHM}} = 30'$ . This is motivated by the effective noise and beam expected for the *LiteBIRD* experiment [12].

The QE approach would, in principle, leave us to deal with many systematics. In practice, to understand the most relevant ones it is useful to use rough estimates of the expected contamination sources and then deal with those that are most prominent. A potentially large source of CMB contamination is a temperature to polarization (T to P) leakage caused by a differential gain systematic. Since the CMB temperature anisotropy signal is 3-4 orders of magnitude larger than primordial B-mode signal, even a small leakage can induce large B-mode power. Therefore, this systematic may be a large hindrance for primordial B-mode studies. In this work we will refine aspects of analysis presented in [7], with our detailed scrutiny limited to focusing on this systematic.

We reiterate that, while all tasks necessary for controlling and understanding this instrument systematic will be performed, it is essential that these efforts be complemented with refined analysis methods that allow mitigation of the such systematics in the observed maps. At the very least, these methods will serve as important null tests, which will need to be performed to claim a robust primordial B-mode signal.

This paper is organized as follows: we begin with a review of different map level instrument systematics in section 2, indicating the levels of contamination that may be induced by different types of distortion fields. We then discuss the details of the respective QE in section 3. The iterative cleaning process which we employ in this work is presented in section 4. Here we also introduce a semi-analytical forecasting procedure that allows us to predict the

expected cleaning of B-mode maps. In section 5 we present details of the simulation where the differential gain systematic is injected and discuss why realistic scan strategy is needed to give credible results. The results from QE analysis on simulated data are presented in section 6. We do this for the ideal case to show the limits of the QE method and for the more realistic, non-ideal case. We further quantify the results in section 7 where we discuss the impact on cleaning on the inferred tensor to scalar ratio  $r$  using a likelihood based approach.

## 2 Distortions of the CMB

In an ideal setting the true CMB polarisation signal would be isolated and easily measured without introducing any distortions. However, in practice the measurement are subject to a number of measurement artefacts which need to be controlled and corrected post measurement via some modelling. These contaminants can be typically characterised by their spin dependence and, as such, readily written in to a set of distortion fields.

Both [7] and [13] use a Müller matrix approach [14] to describe the various systematic and cosmological signals that could affect measurements of the CMB polarization using a series of distortion fields. These distortions can be written as a coupling between different spin combinations of the instrument and observable fields,

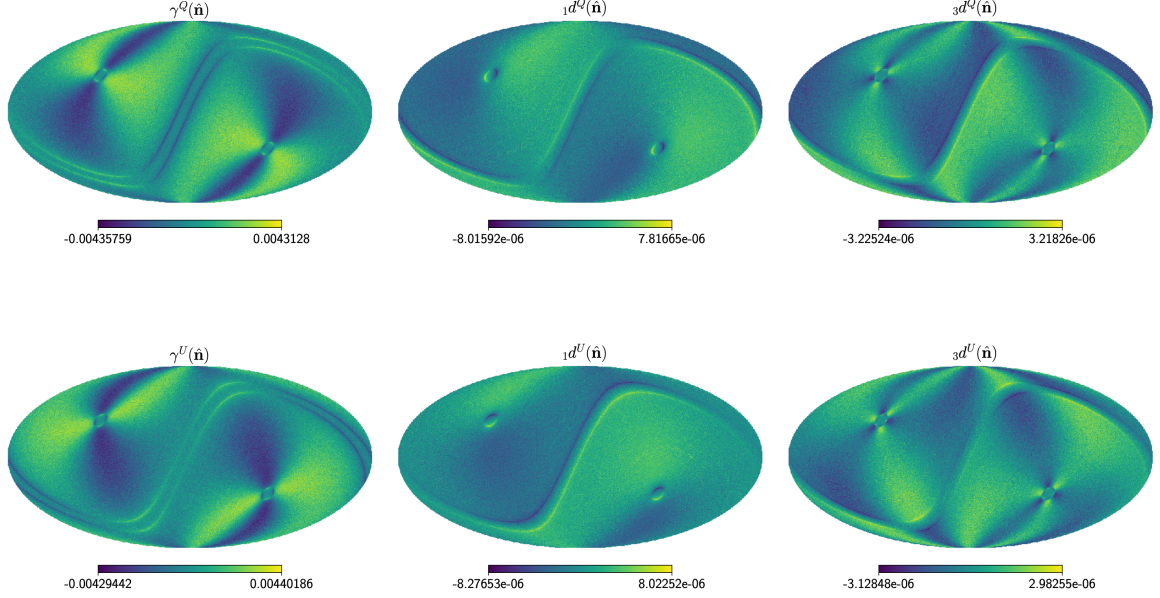
$$\begin{aligned} \pm\delta X(\hat{\mathbf{n}}) = & [a \pm i2\omega](\hat{\mathbf{n}})_{\pm}\tilde{X}(\hat{\mathbf{n}}) + \pm f(\hat{\mathbf{n}})_{\mp}\tilde{X}(\hat{\mathbf{n}}) + \pm\gamma(\hat{\mathbf{n}})\tilde{T}(\hat{\mathbf{n}}) + \sigma_{+1}\mathbf{p}(\hat{\mathbf{n}}) \cdot \mp\tilde{\partial}_{\pm}\tilde{X}(\hat{\mathbf{n}}; \sigma_{\text{fwhm}}) \\ & + \sigma_{\pm}d(\hat{\mathbf{n}})_{\pm}\tilde{\partial}\tilde{T}(\hat{\mathbf{n}}; \sigma_{\text{fwhm}}) + \sigma^2q(\hat{\mathbf{n}})_{\pm}\tilde{\partial}^2\tilde{T}(\hat{\mathbf{n}}; \sigma_{\text{fwhm}}) + \dots \end{aligned} \quad (2.1)$$

where  $\pm\tilde{X}(\hat{\mathbf{n}}) = \tilde{Q} \pm i\tilde{U}$  is the spin  $\pm 2$ , uncontaminated cosmological polarisation signal,  $\tilde{T}$  represents the cosmological temperature signal, and  $\pm\delta X(\hat{\mathbf{n}})$  denotes the total induced distortion, where  $\pm\tilde{\partial}$  denote the spin raising/lowering operators respectively. The  $a(\hat{\mathbf{n}})$  and  $\omega(\hat{\mathbf{n}})$  terms are scalar fields describing an amplitude scaling and a polarisation plane rotation respectively, and  $\pm f(\hat{\mathbf{n}})$  is a spin  $\pm 4$  field which couples the conjugate polarisation fields. The  $\pm\gamma(\hat{\mathbf{n}})$  field is spin  $\pm 2$  field and couples the temperature to polarisation directly.  $\pm\mathbf{p}$  is a spin  $\pm 1$  deflection field that describes direction changes of the photons,  $\pm d(\hat{\mathbf{n}})$  is a spin  $\pm 1$  field and  $q(\hat{\mathbf{n}})$  is a scalar field that couple the first and second derivatives of the temperature field to the polarisation respectively. Since  $\pm X(\hat{\mathbf{n}})$  is a spin  $\pm 2$  field the distortions to this must also be spin  $\pm 2$ .

Equation 2.1 has been constructed such that the top line corresponds to mixing between polarisation and temperature in a known direction on the sky. The bottom line presents terms which involve mixing in a local region of the sky with some directional dependence such that they leak the derivative of the CMB fields, such as a pointing error or lensing. The length scale  $\sigma_{\text{fwhm}}$  corresponds to the width of a Gaussian beam that is smoothing the CMB fields. The terms in the second line are sourced by a simple first order Taylor expansion of the CMB fields around  $\hat{\mathbf{n}}$ .

There are no known processes that cause T to P conversion along the line of sight. Therefore, each of the  $\tilde{T}(\hat{\mathbf{n}})$  containing terms can be attributed to some systematic. Some example distortion fields are depicted in figure 1. Here it is important to note that the form of these distortion fields depends on the scanning strategy and therefore a realistic scan should be employed when assessing the importance of different systematics. We will revisit this detail in section 5.5.

Note that we only consider CMB fields and their distortions in this work. In particular, we do not include foreground fields; we expect that foregrounds would be removed using



**Figure 1.** The left column shows maps of the spin-2  $(\gamma^Q + i\gamma^U)(\hat{n})$  distortion field sourced by a 1% differential gain systematic which causes a T leakage into the polarisation. The second and third columns show the spin-1  $(1d_1 + i_1d_2)(\hat{n})$  and spin-3  $(3d_1 + i_3d_2)(\hat{n})$  distortion fields sourced by a 0.1' differential pointing systematic which causes a leakage of the derivative of the temperature into the polarisation. These levels of gain and pointing are typical of those found in recent CMB experiments [e.g. 15, 16].

standard techniques [17, 18], and we leave for future work any complications arising from interactions between the two methods.

In this paper we will focus on the T to P leakage mediated by the field  $\pm\gamma(\hat{n})$ , which is sourced by differential gain variations in the detector [15, 16] coupled with the instrument scan. Earlier work noted that the  $\pm\gamma(\hat{n})$  field poses the largest potential obstruction to the robust primordial B-mode recovery [7]. Specific details on inclusion of these systematics in the simulated CMB maps will be presented in section 5.

## 2.1 Additional distortion terms

There exist other systematics that contribute at leading order to the distortions described in Equation (2.1). These have been ignored in previous literature [7]. For completeness, here we briefly discuss these ‘new’ terms; these are encoded in the following expression,

$$\begin{aligned} \pm\delta X^{\text{New}}(\hat{n}) = & \sigma_{-1}\mathbf{p}(\hat{n}) \cdot \pm\tilde{\partial}_{\pm}\tilde{X}(\hat{n};\sigma) + \sigma_3\mathbf{p}(\hat{n}) \cdot \pm\tilde{\partial}_{\mp}\tilde{X}(\hat{n};\sigma) + \sigma_5\mathbf{p}(\hat{n}) \cdot \mp\tilde{\partial}_{\mp}\tilde{X}(\hat{n};\sigma) \\ & + \sigma_3\mathbf{d}(\hat{n})_{\mp}\tilde{\partial}\tilde{T}(\hat{n};\sigma) + \sigma^2_4\mathbf{q}(\hat{n})_{\mp}\tilde{\partial}^2\tilde{T}(\hat{n};\sigma). \end{aligned} \quad (2.2)$$

where all the symbols have the same meaning as before.

In particular, note that a differential pointing systematic, which would contribute to the terms containing the  $\pm d(\hat{\mathbf{n}})$  and  ${}_1\mathbf{p}(\hat{\mathbf{n}})$  field in Equation (2.1), also contributes to a number of terms in Equation (2.2). This includes additional coupling of the conjugate of the derivative of the temperature field with the polarisation through the spin-3  ${}_3\mathbf{d}(\hat{\mathbf{n}})$  field. From figure 1, it is clear that this contributes a systematic at a level comparable to the spin-1 contribution,  ${}_1\mathbf{p}(\hat{\mathbf{n}})$ , and as such may not be neglected. In addition the differential pointing will also contribute to the terms containing the fields  ${}_1\mathbf{p}(\hat{\mathbf{n}})$ ,  ${}_5\mathbf{p}(\hat{\mathbf{n}})$ , and  ${}_3\mathbf{p}(\hat{\mathbf{n}})$  which induce polarisation mixing as detailed in [19].

The bias sourced by these additional systematics could in principle hinder a robust claim of primordial  $B$ -mode detection. Neglecting these terms can potentially lead to distortion fields sourced by some systematics remaining undiagnosed, but we leave more detailed explorations of these new systematics to future work.

### 3 Quadratic estimators

For a statistically isotropic (SI) CMB sky, the off-diagonal correlation of the harmonic space covariance matrix  $\langle X_{lm} X'_{l'm'} \rangle \propto \delta_{ll'} \delta_{mm'}$ , for  $X, X' \in [T, E, B]$ . However, secondary anisotropies and measurement artifacts such as the distortion fields described in section 2 can induce off-diagonal correlations. Therefore, by measuring and combining optimally combining these off-diagonal correlations it is possible to draw inferences on the fields that induce deviations from SI. This is commonly referred to as the quadratic estimator (QE) technique and it has been very successfully used to measure the subtle signatures of weak lensing in the CMB and deduce the lensing deflection angle map [6, 20, 21], to test deviations from the standard cosmological assumption of isotropy [22] and also to seek signatures of non-standard physics [23–25]. Some aspects of QE can be discussed quite generally without specific details about the systematics that source the distortions fields, and we refer the reader to [7] for such a discussion. In this work we focus our attention to the QE required to reconstruct the spin-2  $\gamma$  fields which mediates the T to P leakage, the details of which we discuss next.

#### 3.1 Quadratic estimator for the spin-2 $\gamma$ field

We improve on the work presented in [7], by first deriving the full sky QE (i.e without making the flat sky approximation) and then presenting its efficient real space form. We begin by writing the map level model for the observed, distorted CMB sky which is given by the following expression,

$$\pm X(\hat{\mathbf{n}}) = \pm \tilde{X}(\hat{\mathbf{n}}) \star B(\hat{\mathbf{n}}) + \pm \gamma(\hat{\mathbf{n}}) \cdot [\tilde{T}(\hat{\mathbf{n}}) \star B(\hat{\mathbf{n}})] + N(\hat{\mathbf{n}}), \quad (3.1)$$

where,  $\star$  represents a convolution operation and  $\cdot$  represents a scalar multiplication,  $B$  denotes the beam,  $\pm \gamma(\hat{\mathbf{n}})$  represents the T to P leakage fields, and finally  $N(\hat{\mathbf{n}})$  represents the measurement noise. Given that  $\pm X$  and  $\pm \gamma$  are spin two fields, they can be decomposed in the spin weighted spherical harmonic basis as follows,

$$\pm X(\hat{\mathbf{n}}) \equiv \sum_{lm} \pm X_{l_1 m_1} \pm {}_2 Y_{l_1 m_1}(\hat{\mathbf{n}}) = Q(\hat{\mathbf{n}}) \pm iU(\hat{\mathbf{n}}), \quad (3.2a)$$

$$\pm \gamma(\hat{\mathbf{n}}) \equiv \sum_{LM} \pm \gamma_{LM} \pm {}_2 Y_{LM}(\hat{\mathbf{n}}) = \gamma^Q(\hat{\mathbf{n}}) \pm i\gamma^U(\hat{\mathbf{n}}), \quad (3.2b)$$



where the spherical harmonic coefficients can be expressed in terms of the scalar  $E$  and pseudo scalar  $B$  as follows,

$$E_{l_1 m_1} = -\frac{1}{2} (+X_{l_1 m_1} + -X_{l_1 m_1}) ; B_{l_1 m_1} = -\frac{1}{2i} (+X_{l_1 m_1} - -X_{l_1 m_1}) , \quad (3.3a)$$

$$\gamma_{LM}^E = -\frac{1}{2} (+\gamma_{LM} + -\gamma_{LM}) ; \gamma_{LM}^B = -\frac{1}{2i} (+\gamma_{LM} - -\gamma_{LM}) . \quad (3.3b)$$

Given these definition the harmonic space coefficients of expansion of the contaminant spin-2 field is given by,

$$\begin{aligned} \pm \delta X_{l_1 m_1} &= \sum_{LM} \sum_{l_2 m_2} \pm \gamma_{LM} \tilde{T}_{l_2 m_2} \int d\hat{n} \pm_2 Y_{LM}(\hat{n}) Y_{l_2 m_2}(\hat{n}) \pm_2 Y_{l_1 m_1}^*(\hat{n}) , \\ &= \sum_{LM} \sum_{l_2 m_2} \pm \gamma_{LM} \tilde{T}_{l_2 m_2} \pm I_{M m_2 m_1}^{L l_2 l_1} , \end{aligned} \quad (3.4)$$

where that both  $\pm \delta X_{l_1 m_1}$  and  $\tilde{T}_{l_2 m_2}$  fields are beam convolved. We note that the integral  $\pm I_{M m_2 m_1}^{L l_2 l_1}$  has the property:  $+I_{M m_2 m_1}^{L l_2 l_1} = (-1)^\ell - I_{M m_2 m_1}^{L l_2 l_1}$  where  $\ell \equiv L + l_1 + l_2$ . Motivated by this property we define the even and odd parity projection operators as:  $P_{e/o} = \frac{(1 \pm (-1)^\ell)}{2}$ , which as we will see allows to condense a lot of the algebra that follows. Given all the definitions, Equation (3.4) can be re-expressed in the following form,

$$\delta E_{l_1 m_1} = -\frac{1}{2} \sum_{LM} \sum_{l_2 m_2} \left( +\gamma_{LM} \tilde{T}_{l_2 m_2} + I_{M m_2 m_1}^{L l_2 l_1} + -\gamma_{LM} \tilde{T}_{l_2 m_2} - I_{M m_2 m_1}^{L l_2 l_1} \right) , \quad (3.5a)$$

$$\delta B_{l_1 m_1} = \frac{i}{2} \sum_{LM} \sum_{l_2 m_2} \left( +\gamma_{LM} \tilde{T}_{l_2 m_2} + I_{M m_2 m_1}^{L l_2 l_1} - -\gamma_{LM} \tilde{T}_{l_2 m_2} - I_{M m_2 m_1}^{L l_2 l_1} \right) . \quad (3.5b)$$

The above equations can be further reduced to be expressed in terms of the  $\gamma_{LM}^E$  and  $\gamma_{LM}^B$ , which after some simple algebra can be expressed in the following form,

$$\delta E_{l_1 m_1} = \sum_{LM} \sum_{l_2 m_2} \left[ \gamma_{LM}^E \tilde{T}_{l_2 m_2} + I_{M m_2 m_1}^{L l_2 l_1} P_e + i \gamma_{LM}^B \tilde{T}_{l_2 m_2} + I_{M m_2 m_1}^{L l_2 l_1} P_o \right] , \quad (3.6a)$$

$$\delta B_{l_1 m_1} = \sum_{LM} \sum_{l_2 m_2} \left[ \gamma_{LM}^B \tilde{T}_{l_2 m_2} + I_{M m_2 m_1}^{L l_2 l_1} P_e - i \gamma_{LM}^E \tilde{T}_{l_2 m_2} + I_{M m_2 m_1}^{L l_2 l_1} P_o \right] . \quad (3.6b)$$

The polarization contamination in the measured  $E$  or  $B$  fields generated by the  $\gamma_{LM}^B$  and  $\gamma_{LM}^E$  fields can be treated separately by choosing a specific parity. For example,  $\delta B_{l_1 m_1}$  for the  $\ell = \text{even}$  parity is given by,

$$\delta B_{l_1 m_1} = \sum_{LM} \sum_{l_2 m_2} \gamma_{LM}^B \tilde{T}_{l_2 m_2} + I_{M m_2 m_1}^{L l_2 l_1} P_e , \quad (3.7)$$

from which we will derive an estimator that will allow us to reconstruct  $\gamma_{LM}^B$ . Choosing the odd parity mode for  $\delta B$  will allow us to derive a QE that will allow us to reconstruct  $\gamma_{LM}^E$ . Note that QE derivations for other mode combinations follow a near identical procedure. To illustrate the key points we now carry forward the derivation of the QE for  $\gamma_{LM}^B$  starting from Equation (3.7).

The cross correlation between the observed temperature anisotropy map with the observed B-mode maps, under an ensemble average is given by the following expression,

$$\left\langle B_{l_1 m_1}^{\text{obs}} (T_{l'_1 m'_1}^{\text{obs}})^* \right\rangle = \sum_{LM} \gamma_{LM}^B \tilde{C}_{l'_1}^{TT} + I_{M m'_1 m_1}^{L l'_1 l_1} P_e . \quad (3.8)$$

where we have implicitly assumed that the correlation between the true temperature and the true B-mode map is zero owing to parity arguments.<sup>1</sup> Throughout this work  $\tilde{C}_l^{TT}$  represents the beam convolved power spectrum of the primordial CMB temperature signal. We now introduce another identity (see appendix A for details) associated with the integral  $I$ ,

$$\sum_{m_1 m_2} \pm I_{M m_2 m_1}^{L l_2 l_1} \pm I_{M' m_2 m_1}^{L' l_2 l_1} = \frac{(H_{l_2 l_1}^L)^2}{2L+1} \delta_{LL'} \delta_{MM'}, \quad (3.9)$$

where  $H_{l_2 l_1}^L$  is defined in terms of Wigner-3j symbol as,  $H_{l_2 l_1}^L \equiv \sqrt{\frac{(2L+1)(2l_2+1)(2l_1+1)}{4\pi}} \begin{pmatrix} L & l_2 & l_1 \\ -2 & 0 & 2 \end{pmatrix}$ .

Using this identity, the estimator for the correlation in equation (3.8) can be shown to reduce to the following form,

$$\sum_{m_1 m'_1} B_{l_1 m_1}^{\text{obs}} (T_{l'_1 m'_1}^{\text{obs}})^* + I_{M' m'_1 m_1}^{L l'_1 l_1} = \hat{\gamma}_{LM}^B \tilde{C}_{l'_1}^{TT} P_e \frac{(H_{l'_1 l_1}^L)^2}{2L+1}. \quad (3.10)$$

Note that the ensemble average from (3.8) is no longer included here. In reality, we only have access to a single realization of the observed polarization fields when estimating  $\gamma_{LM}^B$  and  $\gamma_{LM}^E$ . This is also why it is necessary to replace the  $\gamma_{LM}^B$  symbol in (3.8) with the symbol for the estimate  $\hat{\gamma}_{LM}^B$  in (3.10). We can easily invert equation (3.10) to construct an estimator for  $\gamma_{LM}^B$  given by,

$$(\hat{\gamma}_{LM}^B)_{l_1 l'_1} = \frac{\sum_{m_1 m'_1} B_{l_1 m_1}^{\text{obs}} (T_{l'_1 m'_1}^{\text{obs}})^* + I_{M' m'_1 m_1}^{L l'_1 l_1}}{F_{l'_1 l_1}^L}, \quad (3.11)$$

where  $F_{l'_1 l_1}^L \equiv \tilde{C}_{l'_1}^{TT} P_e \frac{(H_{l'_1 l_1}^L)^2}{2L+1}$ . This however is only an estimator for  $\gamma_{LM}^B$  from a single multipole pair  $(l, l')$ . It is now possible devise a minimum variance estimator (MVE) by carrying out the inverse variance weighted sum of the estimator across all possible multipole pairs. For this purpose we begin by evaluating the variance of the estimator for a given multipole pair and this is given by the following expression,

$$\begin{aligned} c_{l_1 l'_1}^L &\equiv \langle \hat{\gamma}_{LM}^B (\hat{\gamma}_{L_2 M_2}^B)^* \rangle = \sum_{m_1 m'_1} \sum_{m_2 m'_2} \frac{B_{l_1 m_1}^{\text{obs}} (T_{l'_1 m'_1}^{\text{obs}})^* (B_{l_2 m_2}^{\text{obs}})^* T_{l'_2 m'_2}^{\text{obs}} + I_{M' m'_1 m_1}^{L l'_1 l_1} + I_{M'_2 m'_2 m_2}^{L_2 l'_2 l_2}}{F_{l'_1 l_1}^L F_{l'_2 l_2}^{L_2}}, \\ &= \frac{(2L+1)}{(H_{l'_1 l_1}^L)^2} \frac{\hat{C}_{l_1}^{BB} \hat{C}_{l'_1}^{TT}}{\tilde{C}_{l'_1}^{TT} \tilde{C}_{l'_1}^{TT} (P_e)^2}. \end{aligned} \quad (3.12)$$

The  $\hat{C}_{l_1}^{TT}$  and  $\hat{C}_{l_1}^{BB}$  terms are the power spectra estimated from the observed temperature and B-mode polarization fields respectively

Performing an inverse variance weighted sum of the estimator in (3.11) over  $l_1 l'_1$  yields the MVE QE,

$$\hat{\gamma}_{LM}^B = N_L^{\gamma^B} \sum_{l_1 m_1} \sum_{l'_1 m'_1} \frac{B_{l_1 m_1}^{\text{obs}} (T_{l'_1 m'_1}^{\text{obs}})^* \tilde{C}_{l'_1}^{TT} + I_{M m'_1 m_1}^{L l'_1 l_1} P_e}{\hat{C}_{l_1}^{BB} \hat{C}_{l'_1}^{TT}}, \quad (3.13)$$

<sup>1</sup>For this particular TB QE it is important to note that this estimator does not suffer from any mean field bias and therefore we will not address this detail further.



where  $N_L^{\gamma^B}$  is a normalization, which is also the reconstruction noise (i.e. the power spectrum of the noise in the reconstructed  $\gamma_{LM}^B$  map) which is given by inverse of the variances of all the modes added in parallel, specifically,

$$N_L^{\gamma^B} = \left[ \sum_{l_1 l'_1} \frac{(H_{l'_1 l_1}^L)^2 (\tilde{C}_{l'_1}^{TT} P_e)^2}{(2L+1) \hat{C}_{l_1}^{BB} \hat{C}_{l'_1}^{TT}} \right]^{-1}. \quad (3.14)$$

Note that in Equation (3.13) and Equation (3.14) only the even parity modes (i.e.  $L+l_1+l'_1 \rightarrow \text{Even}$ ) contribute, which only corresponds to only half the elements in the harmonic space covariance matrix. One can show that the other half of the  $TB$  harmonic space covariance matrix encodes information on the E-modes of the spin-2  $\gamma$  field. Following the same procedure as described above, considering the  $\ell = \text{Odd}$  modes, it can be shown that the estimator for  $\gamma_{LM}^E$  is given by the following expression,

$$\hat{\gamma}_{LM}^E = -i N_L^{\gamma^E} \sum_{l_1 m_1} \sum_{l'_1 m'_1} \frac{B_{l_1 m_1}^{\text{obs}} (T_{l'_1 m'_1}^{\text{obs}})^* \tilde{C}_{l'_1}^{TT} + I_{M m'_1 m_1}^{L l'_1 l_1} P_o}{\hat{C}_{l_1}^{BB} \hat{C}_{l'_1}^{TT}}, \quad (3.15)$$

with the reconstruction noise, analogously given by the following expression,

$$N_L^{\gamma^E} = \left[ \sum_{l_1 l'_1} \frac{(H_{l'_1 l_1}^L)^2 (\tilde{C}_{l'_1}^{TT} P_o)^2}{(2L+1) \hat{C}_{l_1}^{BB} \hat{C}_{l'_1}^{TT}} \right]^{-1}. \quad (3.16)$$

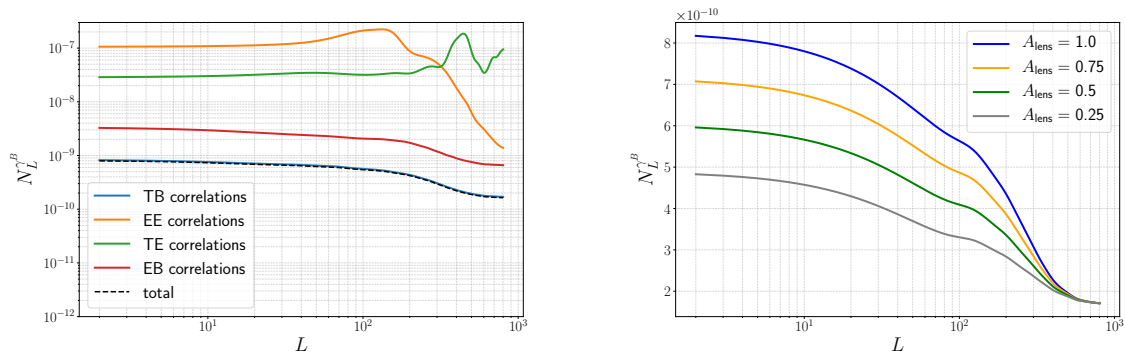
While the forms of the reconstruction noise for  $\hat{\gamma}_{LM}^B$  and  $\hat{\gamma}_{LM}^E$  are nearly the same, they differ in the parity of modes that contribute to the sum and therefore their numerical values are not identical values. These are curved sky equivalents of the flat sky estimators presented in Equation (10) of [7].

### 3.2 The reconstruction noise

When reconstructing the distortion fields, in our case  $\gamma_{LM}^E$  and  $\gamma_{LM}^B$ , the reconstruction noise determines which harmonic modes of these fields can be recovered. Multipoles that are dominated by reconstruction noise cannot be properly reconstructed. Therefore, it is important to perform the reconstruction with as little noise as possible. While the reconstruction noise can be generally reduced by decreasing the measurement noise and increasing the angular resolution of the measurements, we will be interested in minimizing the reconstruction noise for a fixed instrument configuration.

The reconstructions can be performed using QE constructed from a variety of cross correlations:  $EE$ ,  $TE$ ,  $TB$  &  $EB$  and by optimally combining all of them. The reconstruction noise resulting from combining all the estimators naturally is the best. However this requires significant more effort which may not be warranted. This can be assessed by estimating the reconstruction noise for each of the estimators and performing a relative comparison.

We perform this test in three different settings: (i) in the cosmic variance limited ideal case, (ii) in the more realistic, non-ideal case with a noise sensitivity of  $w_{TT}^{-1} = 2.7 \mu\text{K arcmin}$  and a full-width at half-maximum of  $\theta_{\text{FWHM}} = 30'$ , and (iii) assuming the CMBpol configuration with  $w_{TT}^{-1} = 1.4 \mu\text{K arcmin}$  and  $\theta_{\text{FWHM}} = 4'$ . The results from this exercise, for the  $w_{TT}^{-1} = 2.7 \mu\text{K arcmin}$ , and  $\theta_{\text{FWHM}} = 30'$  case study, are summarized in figure 2. In all three cases we find the reconstruction noise associated with the TB QE is the lowest and



**Figure 2.** The left panel shows the reconstruction noise estimated for the respective QE assuming  $A_{\text{lens}} = 1$ . The reconstruction noise for the combination of all the estimators is indicated with a black dashed line. The right panel shows how the TB reconstruction noise varies as a function of  $A_{\text{lens}}$ . Both plots assume instrument noise and beam characteristics compatible with the *LiteBIRD* instrument. Here we display only  $N_L^{\gamma^B}$ , as  $N_L^{\gamma^E}$  shows the same trends.

closely matches the estimated reconstruction noise expected from optimally combining all the different QE. Motivated by this observation we derive all our results using only the TB estimator.

These findings differ from the conclusions drawn in [7] where it is stated that for the CMBpol configurations the EB QE yields the lowest reconstruction noise. This highlights the importance of using detailed case studies for specific distortion fields in order to test the conventional wisdom when applying quadratic estimators. The highest sensitivity choice of correlations for any given distortion field depends on the details of the estimator and it is therefore important to check which choice of correlations yields the most desirable reconstruction noise for each distortion field separately.

### 3.2.1 Impact of de-lensing on reconstruction noise

The B-modes generated by weak lensing of the CMB act as a competing signal for measurement of primordial B-mode signal sourced by tensor perturbations. Therefore many upcoming analysis strategies necessarily include a de-lensing procedure, thereby reducing the additional variance introduced by the lensing signal and consequently improving the measurements or upper bounds on  $r$ . CMB  $B$ -mode power induced by lensing also contributes to the reconstruction noise for the TB correlations. The blind systematic cleaning being proposed in this work can also potentially<sup>2</sup> benefit from the reduced impact of lensing on the observed  $B$ -mode spectrum. Note that in this work we do not carry out the de-lensing procedure, but model it simply by scaling the lensing power spectrum with an amplitude  $A_{\text{lens}}$ .

In the cosmic variance limited case the reconstruction noise scales very simply with different amounts of delensing such that  $N_L^{\gamma} = A_{\text{lens}} N_L^{\gamma}|_{A_{\text{lens}}=1}$ . This simple relationship breaks down in a realistic scenario where the dominant contribution to the  $B$ -mode variance at high  $l$  comes from the instrument noise. This results in de-lensing having little impact on the reconstruction noise for high  $L$  modes. In the more realistic, non-ideal case this translates to modes greater than  $L \sim 600$ , as is seen in figure 2. De-lensing can still reduce the  $B$ -mode

<sup>2</sup>In principle there might be an additional coupling between the lensing and distortion fields which may bias the lensing reconstruction. We leave the exploration of this subtlety to future work.

variance for modes where the instrument noise is not the dominant contribution, and this will be encoded as a reduction in reconstruction noise for lower multipoles.

It is possible to further minimise the impact of the reconstruction noise by carrying out iterative cleaning of the CMB maps and by constructing the optimal filter for  $\hat{\gamma}_{LM}^E$  and  $\hat{\gamma}_{LM}^B$  and we will return to discussing this in section 4.

### 3.3 Efficient real space estimators

These harmonic space estimators derived above involve large sums over multipoles (scaling roughly as  $\sim (l_{\max})^4$ . and also requires evaluation of the Wigner symbols, therefore they are not very computationally efficient. These estimators appear as convolutions in harmonic space and one therefore expects to be able to express them as direct products of some real space fields. This procedure exists and is routinely implemented for the weak lensing estimators [26]. Here we derive an analogous real space operator for reconstructing the spin-2  $\gamma$  field. By using the explicit integral form of  $\pm I_{Mm'l_1}^{LL'l_1}$  we can rewrite Equation (3.13) and Equation (3.15), which after some algebra can be expressed in the following form,

$$\hat{\gamma}_{LM}^B = \frac{N_L^{\gamma^B}}{2} \int d\hat{\mathbf{n}} [ (+_2 A_{B*}(\hat{\mathbf{n}}))^* A_{T*}^{TT}(\hat{\mathbf{n}}) {}_{+2}Y_{LM}(\hat{\mathbf{n}}) + (-_2 A_{B*}(\hat{\mathbf{n}}))^* A_{T*}^{TT}(\hat{\mathbf{n}}) {}_{-2}Y_{LM}(\hat{\mathbf{n}}) ], \quad (3.17)$$

$$\hat{\gamma}_{LM}^E = \frac{-iN_L^{\gamma^E}}{2} \int d\hat{\mathbf{n}} [ (+_2 A_{B*}(\hat{\mathbf{n}}))^* A_{T*}^{TT}(\hat{\mathbf{n}}) {}_{+2}Y_{LM}(\hat{\mathbf{n}}) - (-_2 A_{B*}(\hat{\mathbf{n}}))^* A_{T*}^{TT}(\hat{\mathbf{n}}) {}_{-2}Y_{LM}(\hat{\mathbf{n}}) ]. \quad (3.18)$$

where  $*$  indicates the complex conjugate and the real space fields are defined as follows,

$${}_{\pm 2}A_{B*}(\hat{\mathbf{n}}) = \sum_{l_1 m_1} \frac{1}{\tilde{C}_{l_1}^{BB}} (B_{l_1 m_1}^{\text{obs}})^* {}_{\pm 2}Y_{l_1 m_1}(\hat{\mathbf{n}}), \quad (3.19)$$

$$A_{T*}^{TT}(\hat{\mathbf{n}}) = \sum_{l_1 m_1} \frac{\tilde{C}_{l_1}^{TT}}{\tilde{C}_{l_1}^{TT}} (T_{l_1 m_1}^{\text{obs}'})^* Y_{l_1 m_1}(\hat{\mathbf{n}}), \quad (3.20)$$

where  $X, Y \in [T, E, B]$ . Since these real space fields can be computed independently and merely involve a few spin harmonic transforms, as opposed to explicit multipole sums and evaluations of Wigner symbols, these are significantly more numerically efficient.

## 4 Iterative cleaning

The QE technique detailed in the previous section provides an excellent tool for diagnosis and reconstruction of potential contaminants, which we demonstrate in section 6. However, in this section we shift our attention to discussing how the reconstructed distortion fields can be used to optimally de-contaminate the observed CMB maps. A cleaned CMB map has a lower reconstruction noise, which in effect allows for uncovering the components of the distortions fields that were noise dominated in the original map. These additional components of the reconstructed distortion fields can then be fed back to the cleaning algorithm. This translates to the cleaning of additional contaminated modes. We will refer to this procedure as iterative cleaning; this procedure allows for a more detailed recovery of the distortion fields. We now discuss how the reconstructed distortion fields can be optimally combined with the contaminated maps to yield de-contaminated maps.

#### 4.1 Optimally de-contaminating the CMB maps

The residual contamination in the de-contaminated B-mode map is given by,

$$B_{lm}^{\text{res}} = \delta B_{lm} - \delta \hat{B}_{lm}. \quad (4.1)$$

where  $\delta B$  denotes the true contamination and  $\delta \hat{B}$  is the estimated map of contamination.  $\delta \hat{B}$  sourced by the B-mode component of the  $\gamma$  field can be estimated using the following expression,

$$\delta \hat{B}_{l_1 m_1} = \sum_{LM} \sum_{l_2 m_2} \hat{\gamma}_{LM}^B \tilde{T}_{l_2 m_2} + I_{M m_2 m_1}^{L l_2 l_1} f_{l_2}^L, \quad (4.2)$$

which is the same as Equation (3.6b), except that the distortion field is replaced by the estimated distortion field using the QE as described in section 3 and we have introduced the weights  $f_{l_2}^L$  which need to be determined. We define an optimal cleaning algorithm as one that minimizes the power spectrum of the residual contamination after each iteration of cleaning. Given Equation (4.2), the angular power spectrum of the residual contamination map is given by the following expression,

$$C_{l_1}^{BB, \text{res}} = \sum_{L l_2} \frac{(H_{l_2 l_1}^L)^2}{(2l_2 + 1)} \left[ \tilde{C}_L^{\gamma^B \gamma^B} \tilde{C}_{l_2}^{TT} - 2 \tilde{C}_L^{\gamma^B \gamma^B} \tilde{C}_{l_2}^{TT} f_{l_2}^L + \hat{C}_L^{\gamma^B \gamma^B} \hat{C}_{l_2}^{TT} (f_{l_2}^L)^2 \right], \quad (4.3)$$

and the optimal  $\delta \hat{B}$  can be estimated by solving for the weights  $f_{l_2}^L$  that minimize Equation 4.3. Taking the derivative with respect to  $f_{l_2}^L$  to calculate the minimum of the residual,  $C_l^{BB, \text{res}}$ , results in a filter of the form,

$$f_{l_2}^L = \frac{\tilde{C}_L^{\gamma^B \gamma^B} \tilde{C}_{l_2}^{TT}}{\hat{C}_L^{\gamma^B \gamma^B} \hat{C}_{l_2}^{TT}}. \quad (4.4)$$

It is useful to think of this filter as being composed of two separable parts  $f_{l_2}^L = f_L^{\gamma^B} f_{l_2}^T$ , where

$$f_L^{\gamma^B} = \frac{\tilde{C}_L^{\gamma^B \gamma^B}}{\hat{C}_L^{\gamma^B \gamma^B}}; f_{l_2}^T = \frac{\tilde{C}_{l_2}^{TT}}{\hat{C}_{l_2}^{TT}}, \quad (4.5)$$

which can be understood as being the corresponding Wiener filters for  $\gamma_{LM}^B$  and  $T_{lm}$  fields, on noting that  $\hat{C}_l = \tilde{C}_l + N_l$ . A similar calculation can be carried through for estimating the contamination sourced by the E-mode component of the  $\gamma$  field. This parallels closely the algorithm followed in de-lensing of the CMB sky [6].

Given the Wiener filtered maps  $\hat{\gamma}^{E, \text{WF}}$ ,  $\hat{\gamma}^{B, \text{WF}}$  and  $T^{\text{WF}}$ , the decontaminated polarization maps are given by the following estimator,

$$\pm X^{\text{clean}, i}(\hat{\mathbf{n}}) = \pm X^i(\hat{\mathbf{n}}) - \pm \hat{\gamma}^{\text{WF}, i}(\hat{\mathbf{n}}) T^{\text{WF}, i}(\hat{\mathbf{n}}), \quad (4.6)$$

where we have again used the intergral form of  $\pm I$  to express the multipole sum in Equation 4.2 in its equivalent and efficient real space form. Note that throughout this derivation we work with the beam convolved fields.

The index ‘i’ in the above equation indicates the cleaning iteration. For each iteration of the cleaning beyond the zeroth, the cleaned polarization fields from the previous iteration become the new observed fields to be passed to the QE as well as the cleaning estimator.

As expected, the temperature field remains unaltered through this cleaning process. Note that since the reconstructed  $\gamma$  fields and the corresponding reconstruction noise estimates are continuously updated, the Wiener filters must be freshly estimated at each iteration. This cleaning process is repeated until the reconstruction noise and the B-mode power spectrum converges to their respective floors. We reiterate that in this case study we focus on the details of the iterative cleaning algorithm for the T to P leakage distortion sourced by differential gain, however, this can be generalized to the full range of distortions described in section 2.

#### 4.1.1 Gaussian filters

While it is important that the filters lead to the smallest residual contamination after each iteration of the cleaning, it is also important that the filter prevents the cleaning process from introducing excess bias. While we have shown that the Wiener filters are the optimal filters that minimize  $C_l^{BB, \text{res}}$ , in our numerical experiments working with idealized low noise simulations we find that Wiener filters tend to overestimate the contamination for modes where the reconstruction noise is high, thereby making the iterative procedure have an undesirable non-convergent behaviour. We understand this to be a special feature of a T to P leakage systematic in which  $T \gg B$  and therefore even a small error in the reconstructed  $\gamma$  maps can lead to a large errors in the de-contaminated the B-mode maps in particular. To prevent this we propose a Gaussian filtering scheme,

$$f_L^\gamma = A \exp \left( - \left[ \frac{\hat{C}_L^\gamma}{\hat{C}_L^\gamma - N_L^\gamma} \right]^2 \right), \quad (4.7)$$

where the normalization  $A$  is set such that  $\max(f_l^\gamma) = 1$ . In our numerical simulations we perform, we find this to be a convergent scheme in all cases (unlike Wiener filtering), as it is more aggressive in suppressing modes that are contaminated by noise, thereby preventing excess bias from being introduced into the cleaning.

#### 4.2 Forecasting the reconstruction noise and $C_l^{BB}$ floors

Due to the imperfect reconstruction of the  $\gamma$  fields, it is in practice not possible to perfectly decontaminate the polarization maps using this procedure. To answer this question we have devised a forecasting procedure that enables us to predict the reconstruction noise and  $C_l^{BB}$  floors that the iterative cleaning procedure should in principle achieve.

Making these forecasts involves evaluating the following algorithm. We begin by making an estimate of the reconstruction noise under the assumption that polarization map can be perfectly cleaned. After this initialization we iterate over the following steps until convergence:

- Use the estimated reconstruction noise to simulate Weiner filtered  $\gamma$  maps, using the true systematic maps as input.
- Use the filtered  $\gamma$  maps to perform cleaning on a simulation of contaminated polarization maps using the same procedure prescribed in section 4.1.
- Use the mock cleaned maps to make revised estimates of the reconstruction noise.

The  $B$ -mode calculated from the mock cleaned maps, and the estimates of the reconstruction noise were found to converge after five iterations of the above process. This procedure provides a forecast for both the reconstruction noise floor as well as the cleaned power spectrum

characterizing the polarization maps. We can compare these estimates to the reconstruction noise and the polarization power spectra derived from employing the iterative cleaning procedure to assess if the blind cleaning is performing as expected.

Note that this procedure uses information from the true CMB and systematics maps and as such is only useful for testing the analysis pipeline. For an actual experiment, where we can assume no prior knowledge of the contaminants, we will not have the liberty of carrying out such validation tests. For actual data analysis we would carry out iterations until the reconstruction noise converges, as we will demonstrate in section 6.

## 5 Simulating temperature to polarization leakage

### 5.1 Systematic-differential gain

In upcoming experiments, both satellite- and ground-based, control of T to P leakage will be essential. The relative amplitudes of the signals means that even temperature leakage at the percent level could be a significant contaminant to the B-mode signal. This section will describe the simulation of T to P leakage and the scan strategy we consider for our differential gain case study, following the approach of [19, 27]; see those works for a more exhaustive discussion. We reiterate that the quadratic estimator approach is applicable to a wide range of distortions, we are just choosing this systematic for our detailed case study as it generates the  $\gamma$  distortion field that was found to be most important in [7].

The signal observed by a single detector contains both the temperature and modulated polarisation signals and may be written as

$$d^X = (1 + \delta g^X) \tilde{T}(\hat{\mathbf{n}}) + \tilde{Q}(\hat{\mathbf{n}}) \cos(2\psi) + \tilde{U}(\hat{\mathbf{n}}) \sin(2\psi), \quad (5.1)$$

where  $X$  denotes which detector is being considered, and  $\psi$  is the crossing angle (the orientation of the focal plane on the sky). We have represented a possible constant gain or calibration factor by  $\delta g^X$ , but for our case study we only apply this factor to the temperature signal. A calibration or gain difference can also cause other effects, such as an amplification of the polarisation signal, but we focus on this as it is the most significant problem and is the one that can manifest as a  $\gamma$  type distortion.

We consider a pair-differencing experiment, consisting of co-located detector pairs that are oriented 90 degrees apart. whose observed signals are differenced. Ideally the temperature signal would be completely removed by this procedure, however, any mismatch in the detector gain  $\delta g^X$  between the two detectors will result in leakage of the temperature signal into the polarisation map. We write the differenced signal for a detector pair  $i$  as

$$\begin{aligned} d_i &= \frac{1}{2} (d_i^A - d_i^B) \\ &= \frac{1}{2} \left[ (\delta g_i^A - \delta g_i^B) \tilde{T}(\hat{\mathbf{n}}) + 2\tilde{Q}(\hat{\mathbf{n}}) \cos(2\psi) + 2\tilde{U}(\hat{\mathbf{n}}) \sin(2\psi) \right], \end{aligned} \quad (5.2)$$

where  $A$  and  $B$  denote each detector in the co-located pair. The temperature leakage will occur if  $\delta g_i = \delta g_i^A - \delta g_i^B \neq 0$ , and is given by

$$\delta d_i^g = \frac{\delta g_i}{2} \tilde{T}(\hat{\mathbf{n}}), \quad (5.3)$$



which is a spin-0 quantity. This will combine with the spin-2 part of the scan strategy to create a spurious spin-2 signal that contaminates the polarisation measurement. We may describe a given scanning strategy by a real space field (again, see [19, 27] for more details)

$$\tilde{h}_k = \frac{1}{N_{\text{hits}}(\theta, \phi)} \sum_j e^{ik\psi_j(\theta, \phi)}, \quad (5.4)$$

where  $\psi_j$  is the  $j$ th crossing angle of a given pixel, and  $N_{\text{hits}}$  is the total number of measurements in that pixel. The survey mask is described by  $\tilde{h}_0$ , while  $\tilde{h}_2$  and  $\tilde{h}_4$  naturally appear in simple map-making, and various  $k$  values contribute to different systematic effects (see e.g. [19, 27, 28]). In the differential gain case considered here, the systematic couples to  $\tilde{h}_2$  [19, 27], so the spurious signal due to a detector pair  $i$  is

$${}_2(\delta d_i^g) = \frac{1}{2} \tilde{h}_2(\hat{\mathbf{n}})(\delta g_i) \tilde{T}(\hat{\mathbf{n}}). \quad (5.5)$$

In particular for our simulations, we will use a focal plane with two such pairs of detectors, oriented at 45 degrees to one another to allow simultaneous measurement of both  $\tilde{Q}$  and  $\tilde{U}$  signals. The combined systematic contribution from the two detector pairs is given by

$${}_2(\delta d^g) = \frac{1}{2} \tilde{h}_2(\hat{\mathbf{n}})(\delta g_1 - i\delta g_2) \tilde{T}(\hat{\mathbf{n}}) \quad (5.6)$$

where the factor of  $i$  in the second detector term is due to the rotated orientation of 45 degrees.

For simplicity we choose a setup where each pair of the detectors experiences the same differential gain,  $\delta g_1 = \delta g_2 = 10^{-2}$ , which corresponds to a 1% differential gain. This simplification will not affect the generality of the QE results presented, but one may expect slightly different levels of gain mismatch for different focal plane setups.<sup>3</sup>

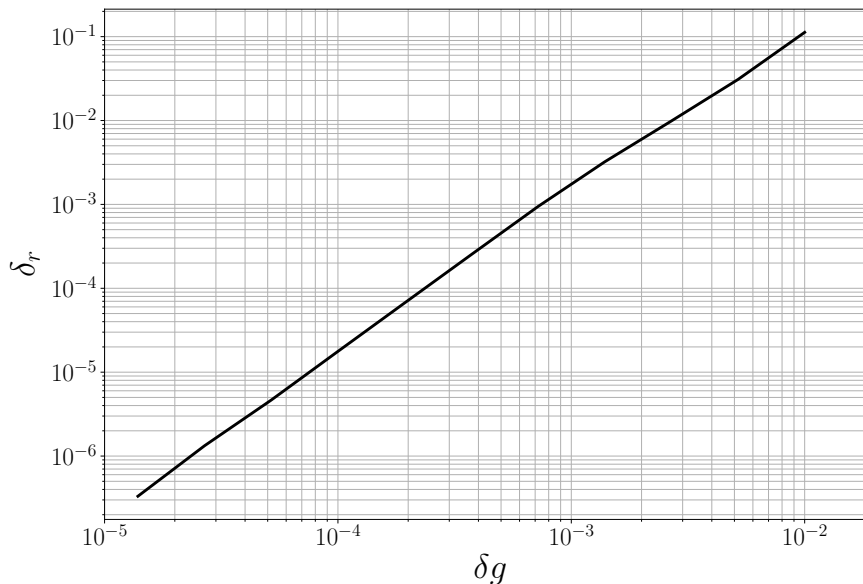
By comparing equation (2.1) to (5.6), we can see that the differential gain be related to the  $\gamma$  distortion as

$$(\gamma^Q \pm i\gamma^U)(\hat{\mathbf{n}})\tilde{T}(\hat{\mathbf{n}}) = \frac{1}{2} \tilde{h}_{\pm 2}(\hat{\mathbf{n}})(\delta g_1 \mp i\delta g_2) \tilde{T}(\hat{\mathbf{n}}). \quad (5.7)$$

## 5.2 Differential gain $r$ -bias

We examine the expected biases on  $r$  that correspond to a range of levels of differential gain. We define the bias sourced by the systematic,  $\delta_r$ , as the difference between the mean of the posterior of the contaminated and uncontaminated spectra. In figure 3 we show the bias on  $r$ ,  $\delta_r$ , for differential gains in the range  $\delta g = 10^{-5} \rightarrow 10^{-2}$ , in the presence of a noise sensitivity of  $w_{TT}^{-1} = 2.7 \mu\text{K arcmin}$  and a full-width at half-maximum of  $\theta_{\text{FWHM}} = 30'$ . We calculate the biases for a fiducial tensor-to-scalar ratio of  $r = 10^{-3}$ . For  $\delta g \lesssim 7 \times 10^{-4}$  the bias is smaller than the  $1\sigma$  statistical variance on  $r$  and cleaning would not be necessary. For values greater than this the bias becomes significant, increasing to  $\sim 100$  times the fiducial  $r$  value for  $\delta g \sim 10^{-2}$ . In this range the bias will have a significant impact on the robustness

<sup>3</sup>There are other methods for handling differential systematics [e.g. 29, 30], however here we study the QE approach in detail rather than performing a comparison of methods. One alternative method being investigated is rotating half wave plates, in which case it isn't clear whether detector differencing should be used [31, 32]. We note that the important systematics in a HWP setup are likely to be different, and it is unclear to what extent our distortion field setup will capture the important systematics. We leave the investigation of the utility of QE in such a situation to future work.



**Figure 3.** The bias on a fiducial tensor-to-scalar ratio of  $r = 10^{-3}$  for levels of differential gain of  $\delta g = 10^{-6} \rightarrow 10^{-2}$  in the presence of a white noise level of  $w_{TT}^{-1} = 2.7 \mu\text{K arcmin}$  and a full-width at half-maximum of  $\theta_{\text{FWHM}} = 30'$ . For values of  $\delta g \sim 1.4 \times 10^{-5}$  the bias is  $\delta_r = 0$  to machine precision.

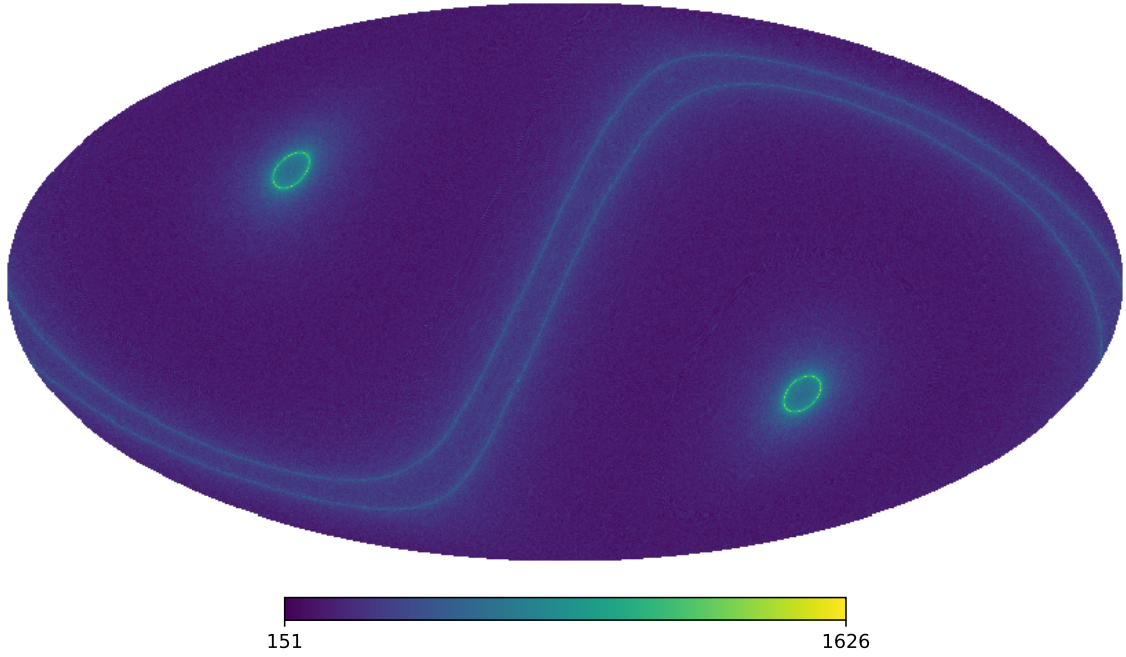
of attempts to measure  $r$ . While the QE cleaning process we present here is able to remove the bias for a range of levels of differential gain, we present results for  $\delta g = 10^{-3}$  in order to demonstrate that it is possible to remove even very large levels of contamination sourced by differential detector gain. Note that this larger level of differential detector gain is typical for a number of contemporary ground-based CMB experiments [e.g. 15, 16], for which we also expect the QE approach to be valuable.

### 5.3 Scan strategy

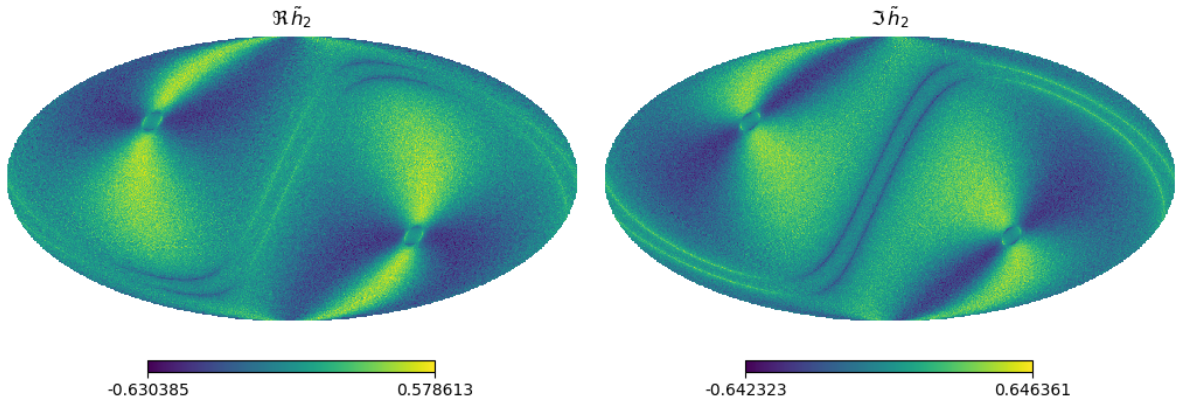
We choose to adopt the Experimental Probe of Inflationary Cosmology (*EPIC*) satellite scan strategy [28]. This will be representative of other future CMB satellite surveys. The design of the EPIC scan strategy optimises crossing angle coverage and is defined by its boresight angle ( $50^\circ$ ), precession angle ( $45^\circ$ ), spin period (1 min), and precession period (3 hrs) (for further details see [27] and [33]). This scan is represented as a list of hits, i.e. datapoints, where each hit is specified by its location on the sky (RA and Dec) and parallactic angle ( $\psi$ ). We expect the QE technique to be equally useful for ground based CMB surveys however, because of the role of the scan strategy in the simulations, we have used a satellite survey in this work for two reasons. Firstly, because no ground-based survey covers the whole sky, and we wanted to avoid complications due to partial sky coverage. Secondly, because ground based surveys differ more between experiments and cannot be simply described by a few parameters as satellite scans can, since they depend on complicated constraints and detailed scheduling choices.<sup>4</sup>

The scan strategy of *EPIC* provides a relatively uniform distribution of hits and crossing angles which should reduce scan coupled differential systematics fairly well. The galactic poles

<sup>4</sup>Although see [34] for some simple approaches to approximate ground-based scan strategies that capture the features relevant for studies such as this one.



**Figure 4.** The hit map of the *EPIC* scan strategy. This survey has been designed to maximise crossing angle coverage, and the hit map is well filled across the full sky with many observations at different orientations. Note that the galactic poles have been prioritised to aid the understanding of foregrounds and galactic science goals hence the higher number density in the hit map in those regimes



**Figure 5.** The real (left panel) and imaginary (right panel) parts of the  $\tilde{h}_2$  field (defined in equation (5.4)), which encodes the spin-2 part of the scan that turns the gain systematic into a spurious spin-2 signal. The  $\tilde{h}_2$  field is dominated by its large scale features, which will result in a low  $l$  dominated systematic.

are observed more frequently to aid in foreground analysis and galactic science goals, which results in some structure appearing in the survey fields as seen in figure 4. The  $\tilde{h}_{\pm 2}$  field (equation (5.4)) encodes the spin-2 part of the scan that turns the gain systematic into a

Simulation Inputs	
Cosmology	$H_0 = 67.4$
	$\Omega_b h^2 = 0.022$
	$\Omega_c h^2 = 0.120$
	$\tau = 0.06$
	$n_s = 0.97$
	$10^9 A_s = 2.2$
	$r = 0.001$
Map-making	$N_{\text{side}} = 2048$
	$\ell_{\text{max}} = 3000$
	$ \delta g_1  = 0.01$
	$ \delta g_2  = 0.01$

**Table 1.** The fiducial cosmological parameters and the map-making inputs for the TOD simulations. The simulated maps are smoothed by a Gaussian beam and noise is added per pixel.

spurious spin-2 signal. This field is shown in figure 5 for the *EPIC* scan strategy. The  $\tilde{h}_2$  is dominated by its large scale features, and this will result in a low  $l$  dominated systematic.

#### 5.4 Simulation

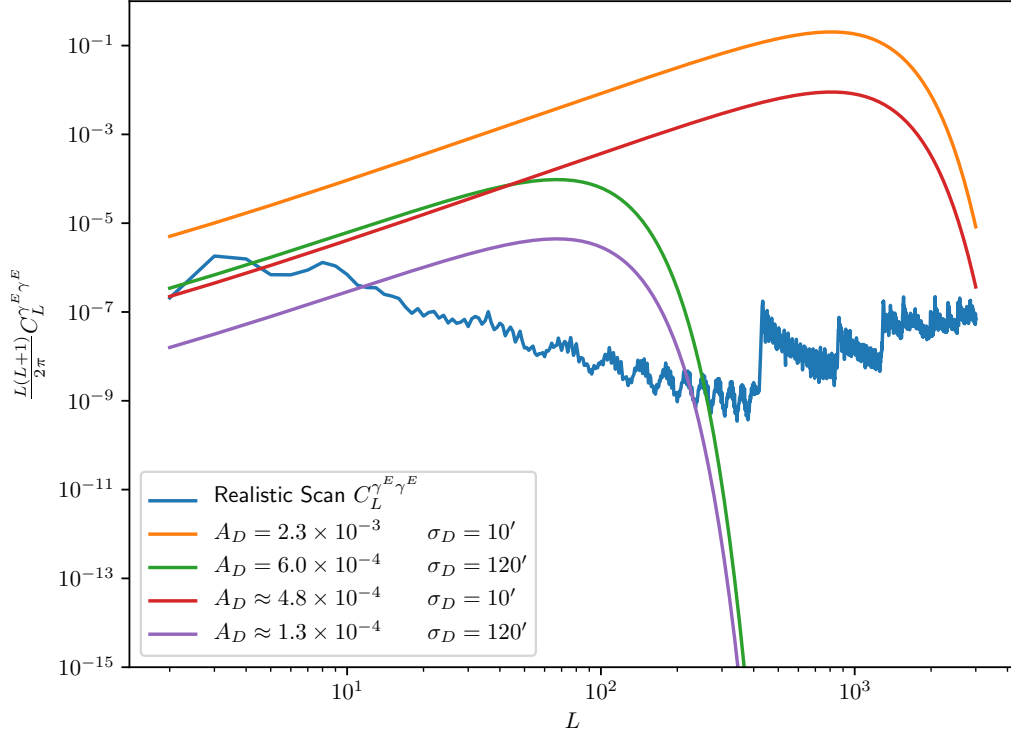
We use a modified version of the code used in [27]. The input to the time ordered data (TOD) simulation code consists of maps of the CMB  $\tilde{T}$ ,  $\tilde{Q}$  and  $\tilde{U}$  fields which are generated using the SYNFAST routine of the HEALPIX package [35]. The input CMB power spectra were created in CAMB using a six parameter  $\Lambda$ CDM cosmological model, specified in table 1 [36]. For the simulations including noise, we include a Gaussian beam post process. A white noise is applied to the data at map level, where a noise is added to each pixel of the level  $w_{TT}^{-1} = 2.7 \mu\text{K arcmin}$  [37].

The simulation computes values for each of the four detectors for each hit as described in equation (5.1), using HEALPIX interpolation to observe the input sky maps at the appropriate location, and using the corresponding parallactic angle  $\psi$  for that hit, generated from the *EPIC* scan strategy. As described in section 5, the  $\psi$  values are offset by  $90^\circ$  for the two detectors within a pair, and by  $45^\circ$  between the two pairs of detectors. The differential gain systematic is added for each pair of detectors by increasing the signal by a factor  $(1 - \delta g_i)$  in the second detector  $d_i^B$  in each pair, where we use  $|\delta g_1| = |\delta g_2| = 0.01$  for the simulations in this work [38]. This level of systematic is indicative of the differential gain seen in recent CMB surveys [e.g. 15, 16], and corresponds to an  $r$  bias of  $\sim 30\sigma$ .

Maps are made from the time streams according to a simple binning map making technique,

$$\begin{pmatrix} Q \\ U \end{pmatrix} = \begin{pmatrix} \langle \cos^2(2\psi_i) \rangle & \langle \cos(2\psi_i) \sin(2\psi_i) \rangle \\ \langle \sin(2\psi_i) \cos(2\psi_i) \rangle & \langle \sin^2(2\psi_i) \rangle \end{pmatrix}^{-1} \begin{pmatrix} \langle d_i \cos(2\psi_i) \rangle \\ \langle d_i \sin(2\psi_i) \rangle \end{pmatrix}, \quad (5.8)$$

where the angle brackets  $\langle \rangle$  denote an average over the measurements in a pixel, and the  $d_j$  here correspond to the detector measurements (i.e. the sum of the timestreams from the two differenced pairs) each with its associated angle  $\psi_j$ .



**Figure 6.** The E-mode spectrum of the  $\pm\gamma(n)$  distortion fields, the orange and green lines show the spectra plotted for the  $A_D^{\text{Max}}$  amplitudes of [7] for coherence scales of  $10'$  and  $120'$  respectively, the red and purple lines show the spectra plotted for the  $A_D^{\text{Min}}$  amplitudes of [7] for the CMBPol satellite estimate corresponding to  $r=0.005$  for coherence scales of  $10'$  and  $120'$  respectively. The blue line shows the realistic spectrum calculated from the a combination of the systematic and the *EPIC* scan strategy that we adopt for our simulations. It is clear here that the approximation used for the distortion power spectra does not describe the realistic case well at higher multipoles. It overestimates the power at high multipoles in comparison to the realistic case for small coherence scales. Additionally the sharp drop off in power after the peak is not consistent with the realistic spectra as is evident for all sets of curves showing the approximation.

### 5.5 Realistic systematic spectra

In previous studies [e.g. 7, 13] the distortion fields of equation (2.1) have been assumed to be statistically isotropic and Gaussian, and defined by power spectra of the form

$$C_l^{DD} = A_D^2 e^{-l(l+1)\sigma_D^2/8\ln 2}, \quad (5.9)$$

where  $A_D$  represents the root mean squared of the distortion field, and  $\sigma_D$  represents a coherence scale beyond which the distortion power spectrum becomes white noise. We show a comparison between this spectrum and the realistic spectrum from our simulations in figure 6. There are significant differences apparent between the approximation of equation (5.9) for the power spectra describing the distortion fields used in previous literature, and the realistic distortion field that is derived from the more realistic simulation used in our study.

Although the spectrum generated from equation (5.9) results in most of the power being at low multipoles similar to the realistic spectra, it does not capture the high  $l$  nature of the realistic distortion fields. The realistic distortion fields have an initial much sharper drop

off, before levelling out, compared to the more gradual drop off of the approximation. In figure 6 we show that the two extreme coherence scales used in [7] both suffer from the same issue that they accrue too much power at high  $l$  compared to the realistic spectra, hence the results found will be biased by this. In our analysis in section 6 we reconstruct the distortion power spectrum up to  $L = 800$  and, as can be seen in figure 6, the shapes of the realistic and approximate spectra differ significantly for this range of multipoles. This shows the advantage of carrying out detailed case studies on individual distortions.

We note that the smallest  $A_D^{\min} \approx 1.3 \times 10^{-4}$  for  $\sigma_D = 120'$  quoted from [7] corresponds to a gain mismatch of  $6.4 \times 10^{-4}$  (0.064%). This is significantly smaller than a realistic gain mismatch of 0.01 (1%), in the analysis in [7] it was found that the QE technique would still be effective for this small value of the systematic.

## 6 Reconstructing and removing the temperature to polarization leakage

Here we discuss the results of employing the statistical analysis methods developed in section 3 and section 4 to contaminated CMB maps simulated as in section 5. Note that we include lensing effects only at the power spectrum level, implying that the off-diagonal elements sourced by weak lensing are not included in our simulations. This is not expected to influence our inference of the reconstruction and removal of the T to P leakage systematic, owing to different spins associated with the two effects.<sup>5</sup>

In all our analyses we iterate over the following steps until convergence:

- Reconstruct map of systematics given some input  $[T, Q, U]$  maps using the QE algorithm. In the first iteration the inputs correspond to the observed maps, while for the subsequent iterations these correspond to the contamination cleaned maps.
- Clean the input maps using the reconstructed  $\gamma$  maps following the optimal cleaning procedure discussed in section 4.1.

Here we reemphasize that the cleaning analysis is agnostic to details of the particular systematic, as evident by the fact that the cleaning procedure only works with the observed maps as inputs. All QE evaluations required in the blind cleaning process are carried out assuming the parameter settings summarized in table 2. We make forecasts for the reconstruction noise and the  $C_l^{BB}$  spectrum that one expects to recover from the iteratively cleaned maps, following procedures outlined in section 4.2. We use these forecasted power spectra as benchmarks for our blind cleaning analysis.

We present the results of this analysis on two different set of simulations, the ideal case and the more realistic, non-ideal case, in sections 6.1 and 6.2 respectively. A discussion with particular emphasis on the measurement of tensor to scalar ratio  $r$  is presented in section 7.

### 6.1 The cosmic variance limits of the blind cleaning algorithm

Here we discuss the results derived from analyses on simulations which are ideal in the sense that they include no measurement noise & beam smoothing and also do not include any lensing induced B-modes. These simulations allow us to probe the limitations of the blind

---

<sup>5</sup>Note that this is not generally true for other instrument systematics. As an example a differential pointing systematic directly couples to the weak lensing effect and in a analogous study the weak lensing induced correlations cannot be ignored.



Parameter	Ideal	Non-ideal
$l_{\max}$	1400	1400
$L_{\max}$	800	800
$N_{\text{side}}$	1024	1024
$w_{TT}^{-1}$	0	$2.7 \mu\text{K arcmin}$
$\theta_{FWHM}$	0	$30'$
$r$	$10^{-3}$	$10^{-3}$
$A_{\text{lens}}$	0	1

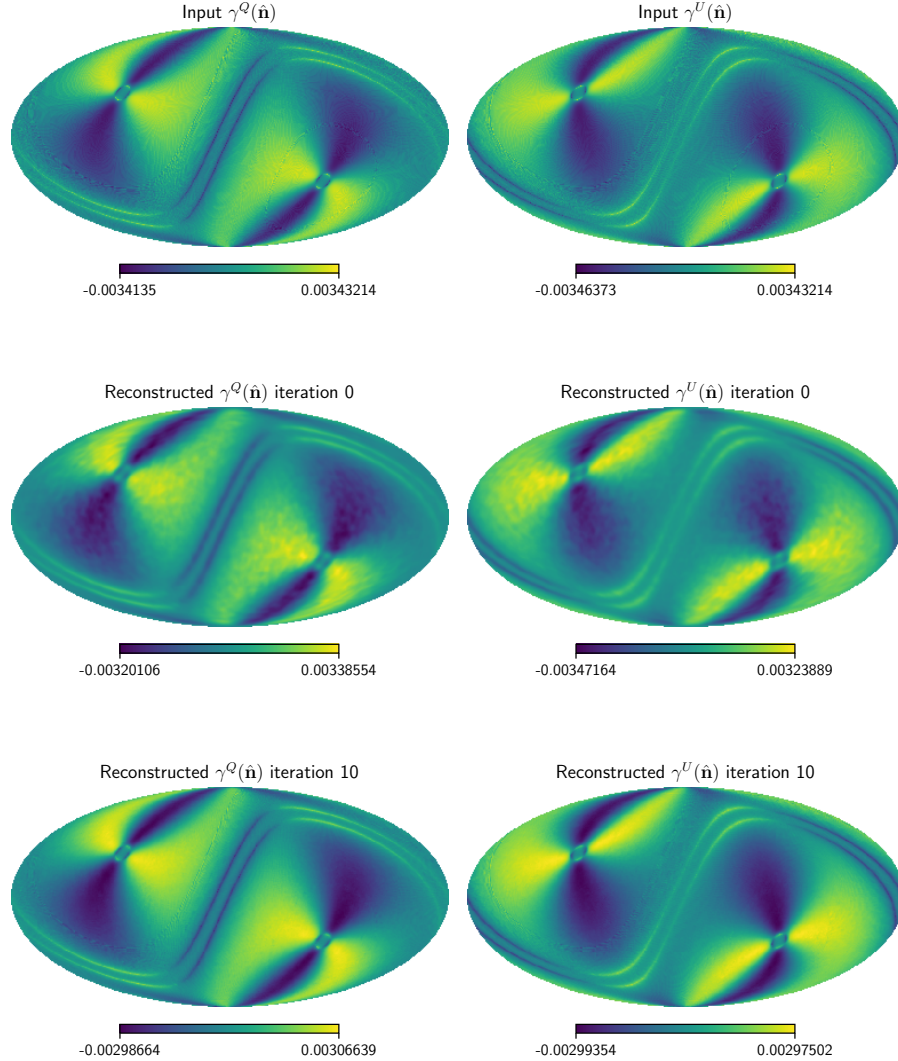
**Table 2.** This table summarizes the QE parameter settings used in all our analyses and also the simulation settings for the two different sets of simulations used in results presented in this section.

cleaning algorithm in this extreme setting, in process highlighting how well this procedure could in principle work.

The reconstructed  $\gamma$  maps at some example iterations of the analysis are depicted in figure 7. Here we note that the systematics reconstructed from the original observed maps are quite noisy as inferred by comparing the top and middle panels of figure 7. This observation is better quantified in figure 8, where by inspecting the reconstruction noise for the 0<sup>th</sup> iteration and comparing it to the true  $\gamma$  power spectrum, it is clear that only  $L \lesssim 30$  multipoles of the  $\gamma$  map can be reliably recovered. The reconstruction noise being high is due to the excess B-mode power sourced by the systematics in the observed CMB maps as seen in figure 9.

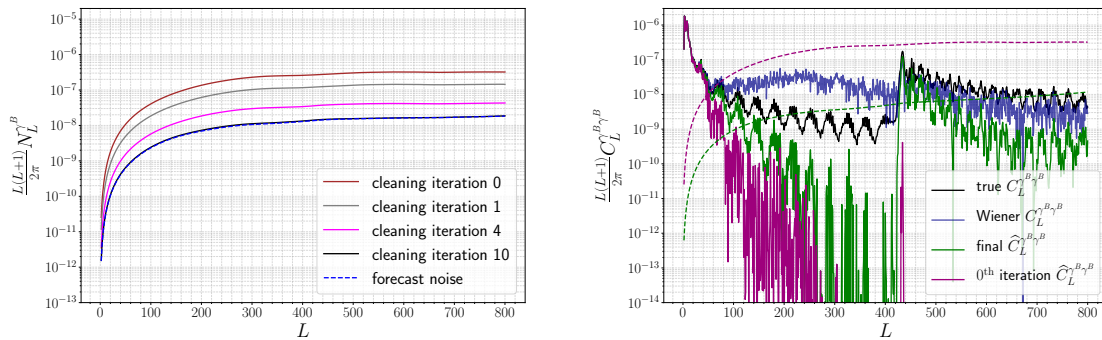
We now use these reconstructed  $\gamma$  maps together with the observed temperature anisotropy map to remove part of the contamination, sourced by modes in the  $\gamma$  maps that have been reliably recovered. This procedure involves using the high SNR modes of the temperature and  $\gamma$  maps, the formal details of which are discussed in section 4.1. In the case of these idealized simulations, we find that the conventional Wiener filtering scheme causes the iterative scheme to diverge after few initial iterations. We suspect this behaviour arises from the fact that the Wiener filter does not sufficiently suppress modes that have a noisy recovery, which combined with the fact that  $T \gg B$ , leads to a faulty cleaning of the polarization maps, in effect adding more power to the B-mode map as opposed to subtracting it. This eventually leads to run away behaviour. We deal with this issue by employing the Gaussian filtering scheme instead (see section 4.1 for details) which mitigates this issue by imposing a stronger suppression of the noisy modes, leading to more stable and convergent results. The right panel of figure 8 depicts the power spectrum of the  $\gamma^B$  maps after the Gaussian and Wiener filters are applied to them, where notably the Wiener filtered maps retain a lot of power from the unreliably recovered modes, as opposed to the Gaussian filtered map where these noisy modes are more heavily suppressed.

After the first cleaning (i.e. cleaning iteration 0), the B-mode power spectrum reduces compared to the spectrum estimated from the observed B-mode map as seen in figure 9. This results in the reconstruction noise of the QE to reduce as can be understood by comparing the curves corresponding to “iteration 0” and “iteration 1” in figure 8. This reduction in the reconstruction noise, facilitates the recovery of modes in the  $\gamma$  map that were dominated by the reconstruction noise in the previous iteration. These newly recovered modes of the systematic map are then fed to the cleaning algorithm to further remove the contamination from the polarization maps. This whole process is repeated until we observe no further improvements in either the reconstruction noise and/or the  $C_\ell^{BB}$  spectrum.



**Figure 7.** This figure depicts the  $\gamma^Q(\hat{n})$  and  $\gamma^U(\hat{n})$  systematic maps that mediate the T to P leakage. The top panels depicts the true systematic maps used to simulate the contaminated CMB maps. The middle panels shows the QE reconstruction of the  $\gamma$  maps from the observed CMB maps for the 0th iteration. The bottom panels depict the reconstructed  $\gamma$  maps after 10 iterations of cleaning and reconstructing of the systematic maps.

On repeating this procedure we see that the contamination in the CMB polarization maps is progressively removed as indicated by the systematic reduction in the amplitude of  $C_l^{BB}$  amplitude in figure 9. Note that initial iterations show relatively big reductions in power, with subsequent iterations resulting in more subtle improvements and the final few iterations show no appreciable updates to the spectrum. The performance of the cleaning process improves if the amplitude of the uncontaminated  $B$ -mode spectrum increases. For the ideal case this translates to improved performance for larger values of  $r$ . We note that even in this perfect setting of no instrument and lensing noise, the recovered  $B$ -mode spectrum is not perfectly cleaned. This can be interpreted as an intrinsic limitation to how well the cleaning

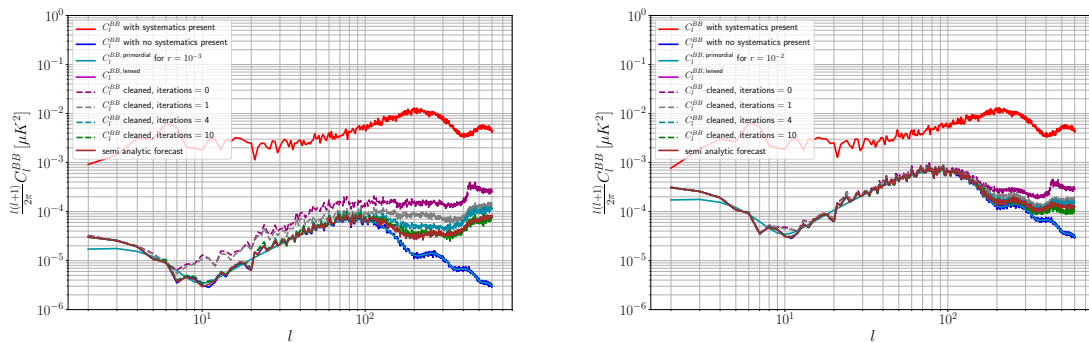


**Figure 8.** The left panel shows the evolution of the reconstruction noise for different iterations of the algorithm. Note that the reconstruction noise reduces with iterations and approaches that predicted using the forecasting procedure. The right panel shows the power spectra of the true and reconstructed and filtered  $\gamma$  maps. The corresponding reconstruction noise curves are plotted for reference. Note that the Wiener filtered,  $\hat{C}_L^{\gamma^B, \text{WF}}$ , spectrum is most reliably recovered for modes where  $C_L^{\gamma^B}$  is much greater than  $N_L^{\gamma^B}$ . Similar results found for  $\gamma^E$ , not show here for brevity.

can in principle perform. Nonetheless, the proposed blind cleaning procedure enables robust removal of contamination power that is roughly two orders of magnitude larger than the injected signal, and yields an unbiased recovery of the true signal at most multipoles.

This systematic reduction in the B-mode power is only possible due to the simultaneously reduction in the reconstruction noise (sourced by reduction in  $C_L^{BB}$ ) as seen in figure 8, which results in robust recovery of the higher multipole of the  $\gamma$  maps (which in turn facilitates more cleaning of the polarization maps). To contrast the effect of iterative cleaning note that while the “iteration 0” only allowed for recovery of the modes  $L \lesssim 30$ , the reconstruction noise associated with the final iteration of cleaning allows robust recovery of modes up to  $L \simeq 800$  as can be seen in right panel of figure 8. This stark improvement in the recovery of high  $L$  modes of the  $\gamma$  maps can be better appreciated by simultaneously comparing the recovered total  $\gamma$  maps shown in the bottom panels of figure 7 to those depicted in the panels above. Note that the total systematic maps is recovered by adding together the filtered maps of systematics estimated at each iteration.<sup>6</sup> The input maps have a higher amplitude than those reconstructed which is primarily a consequence of our maps being filtered and the reconstruction being terminated at  $L_{\text{max}} = 800$ . Here, it is also important to appreciate the non-monotonic nature of the true  $C_L^{\gamma\gamma}$  which features a prominent jump in power at  $L \sim 400$ . This is a consequence of using realistic scan maps in our simulations. There is a corresponding feature in the cleaned B-mode spectra in figure 9. This feature is present because not all modes below  $L \lesssim 800$  are reconstructed, as some intermediate modes which are dominated by reconstruction noise are suppressed. We reiterate that this would not have been observed in studies using the approximate spectra, [e.g. 7, 13], generated using equation (5.9) because of the difference in shape between the realistic and approximate spectra. It is necessary to carry out detailed case studies systematics in order to observe these important details. Unlike in weak lensing studies, for instrument systematics it is not possible to make a generic forecasts as was done in [7].

<sup>6</sup>The  $\gamma$  maps recovered at each iteration do not include the modes that were in effect subtracted from the polarization data in the previous iteration.



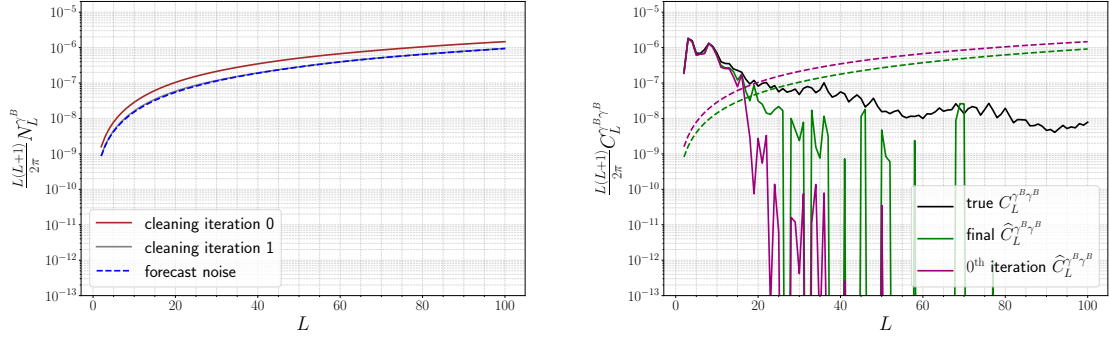
**Figure 9.** This figure depicts the B-mode power spectrum corresponding to the systematic ridden maps, the true cosmological primordial signal and the evolution of the estimated power spectrum across different iterations of the cleaning algorithm. The left and right panels show the cleaning for  $r = 10^{-3}$  and  $r = 10^{-2}$  respectively. Also shown is the prediction for the B-mode power spectrum expected post cleaning evaluated using the forecasting procedure.

Finally, we note that the spectrum converges to the prediction from our forecasting procedure. This is true both for  $C_l^{BB}$  as well as  $N_L^{\gamma\gamma}$  as seen in figure 9 and figure 8 respectively. It is important to note this near consistency for two reasons, (i) it serves as a validation of our blind cleaning algorithm (ii) the actual analysis is performed using the Gaussian filter, while our forecasting procedure continues to use Weiner filters, and the near equivalence of the two solutions suggests that the Gaussian filtering is close to optimal.

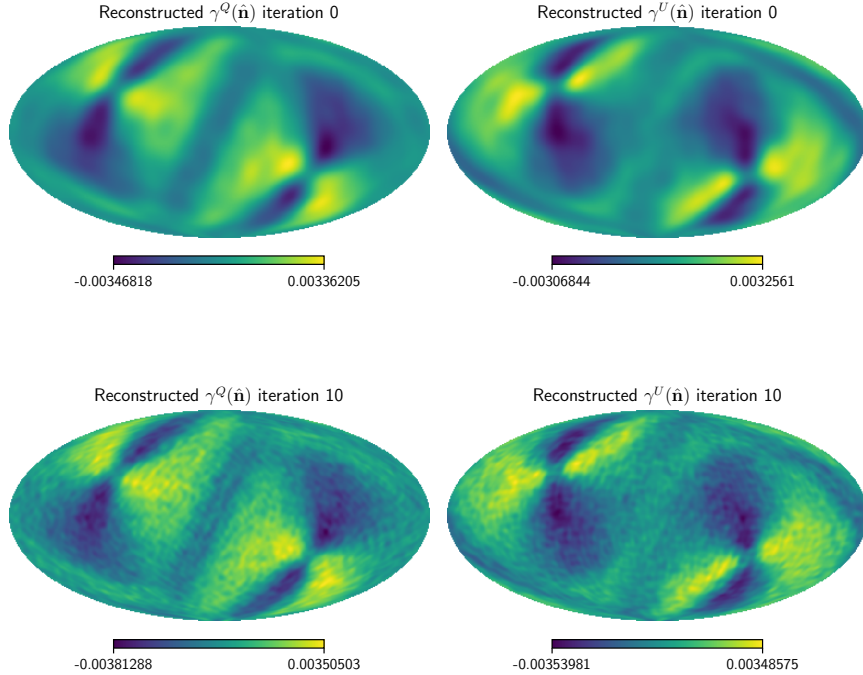
## 6.2 Employing blind systematic cleaning for a non-ideal experiment

Here we discuss results simulations that incorporate weak lensing induced B-modes as well as the measurement noise and beam smoothing in the previously described more realistic, non-ideal case. The assumed measurement noise and beam are summarized in table 2 and correspond to the foreground cleaned Q/U maps that will result from linearly combining the multi-frequency measurements. Unlike in the previous section where these were ignored, here we treat the lensing  $B$ -modes as an important cosmological signal that we recover by carrying out the iterative de-contamination procedure.

We carry out an analysis, identical to that described in the previous section, on these more realistic simulations. The simulations used here primarily differ from those used in the previous section by inclusion of the relatively high noise in the observed maps due to inclusion of lensing and measurement noise. We note that for these relatively high noise simulations, the Wiener filtering schemes is stable and convergent, and the results are very similar to those found when using the Gaussian filtering. We however continue to present results derived from employing the Gaussian filtering scheme through the rest of the paper. The relatively high noise results in a higher QE reconstruction noise floor, which consequently limits the reliable reconstruction of the  $\gamma$  maps to only the large angle modes  $L \lesssim 20$ , even after ten iterations of cleaning as seen in figure 10. This is even reflected in the total reconstructed  $\gamma$  maps as seen in figure 11. However note that the carrying out a number of iterative cleaning procedures does help in recovering some additional features in the reconstructed  $\gamma$  maps, the sharpening of the features in the equatorial plane in the bottom left panel of figure 11 is particularly noticeable. As we will see in section 7 these subtle improvements in recovery of



**Figure 10.** The left panel shows the reconstruction noise  $N_L^{\gamma^B}$  for one iteration of the reconstruction and cleaning process. After a single iteration the cleaning reconstruction noise converges with the forecasted reconstruction noise. The right panel shows  $\hat{C}_L^{\gamma^B \gamma^B}$  for the reconstructed and filtered  $\gamma$  maps before cleaning and after ten iterations of cleaning (green) and the true  $C_L^{\gamma^B \gamma^B}$  is also shown for comparison. Corresponding reconstruction noise shown for reference.

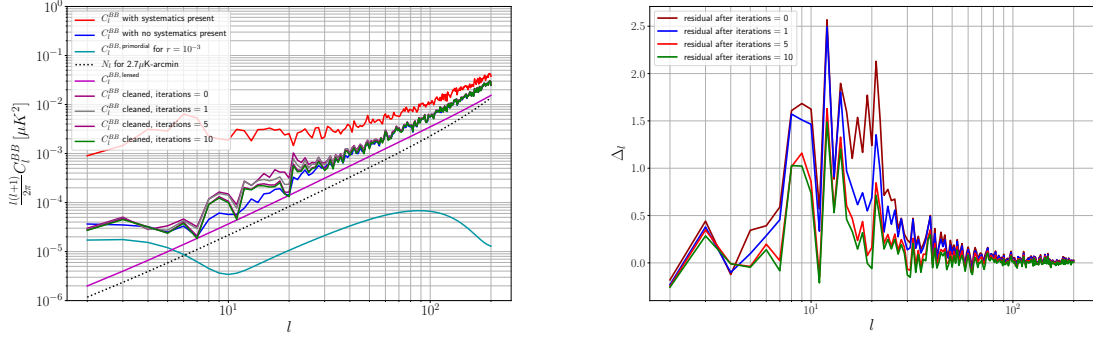


**Figure 11.** The top panels show the reconstructed filtered  $\gamma$  maps recovered from simulated observations, while the bottom panels show those recovered after a few iterations of cleaning. Note that even in the presence of instrument and lensing noise the iterative procedure helps with extracting bits of information on the systematics.

the systematic maps will play a crucial role in more robust removal of contaminations from the observed maps.

We now shift our attention to the evolution of  $C_l^{BB}$  across the cleaning iterations. Carrying out higher iterations of cleaning does make small improvements to the convergence. These subtle but important improvements are highlighted in the right panel of figure 12,





**Figure 12.** The left panel shows cleaned  $B$ -mode power spectra for 1 and 10 iterations of cleaning. The right panel shows the evolution of the relative difference  $\Delta_l = (C_l^{BB, \text{clean}} - C_l^{BB, \text{true}})/C_l^{BB, \text{true}}$  with iterations.

where the relative differences between the spectra derived from the cleaned maps at different iterations and true spectrum are depicted. In section 7 we will highlight the importance of these subtle corrections in the context of measurement of tensor to scalar ratio  $r$ .

## 7 Recovering the tensor-to-scalar ratio

In the previous section we demonstrated that blind systematic cleaning method proposed here can yield nearly un-biased recovery of the true CMB  $B$ -mode power spectrum. Upcoming experiments aim to recover  $r \in [10^{-2}, 10^{-3}]$  [4, 12]. Here we demonstrate that the blind cleaning technique can yield nearly un-biased recovery of  $r$ .

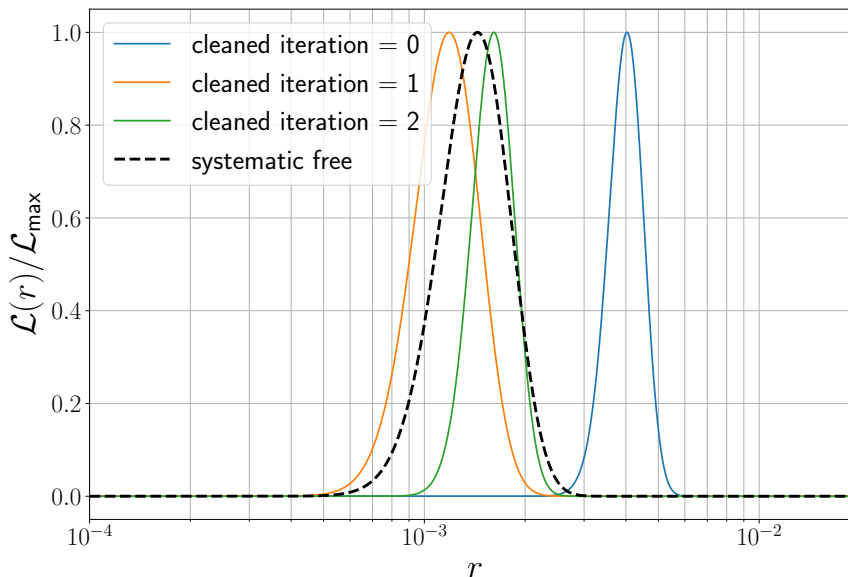
To demonstrate this we carry out a likelihood analysis for which we assume this specific form of the log-likelihood [39, 40], which accounts for the non-Gaussian nature of the power spectrum at low multipoles,

$$\begin{aligned}
 -2 \ln \mathcal{L}(r) = \sum_l (2l+1) & \left[ \frac{\hat{C}_l^{BB}}{r C_l^{BB, GW} + C_l^{BB, \text{lens}} + N_l^{BB}} \right. \\
 & \left. + \ln \left( r C_l^{BB, GW} + A_{\text{lens}} C_l^{BB, \text{lens}} + N_l^{BB} \right) - \frac{2l-1}{2l+1} \ln \left( \hat{C}_l^{BB} \right) \right] + \text{const.},
 \end{aligned} \tag{7.1}$$

where  $\hat{C}_l^{BB}$  denotes the power spectrum estimated from the simulated data, corrected for the instrument beam,  $C_l^{BB, GW}$  is the  $B$ -mode signal generated by primordial gravitational waves evaluated for  $r = 1$ ,  $C_l^{BB, \text{lens}}$  denotes the lensing induced  $B$ -mode spectrum and  $N_l^{BB}$  is the instrument noise power spectrum. We evaluate this likelihood analysis on power spectra derived from the cleaned maps at a number of different iterations and compare the estimated posteriors on  $r$  to those derived from an analysis on a contamination free simulation.

We begin by noting that the  $r$  inferred from the contaminated  $B$ -mode simulations returns a highly biased measurement of  $r \sim 10^{-1}$ , off-set by two orders of from the true value. However, on repeatedly applying the iterative cleaning algorithm to it using either the Wiener or the Gaussian filter, the bias in the measurement of  $r$  reduces. The reduction in  $r$ -bias is largest in the non-ideal setup when using Wiener filtering. As seen in figure 13, after 2 iterations we find that the measured value of  $r$  is consistent with the true value to better than  $1\sigma$ .





**Figure 13.** The evolution of the  $r$ -posterior across different cleaning iterations using the Wiener filter. The bias on  $r$  reduces with each iteration and is remarkably consistent with the posterior derived from the systematic free simulation.

	iteration	$r$	$r_-$	$r_+$
Contaminated	-	113	111	115
Cleaned	0	4.03	3.58	4.53
	1	1.19	0.972	1.5
	2	1.61	1.41	1.88
	3	1.61	1.41	1.88
True	-	1.44	1.16	1.86

**Table 3.** This table presents the central value of  $r$ , and 68% CI upper and lower bounds in units of  $10^{-3}$ . These results assume no de-lensing i.e.  $A_{\text{lens}} = 1$ .

This corresponds to a near perfect removal of a bias of order  $\sim 100$ . Beyond the 2nd iteration of cleaning there is no further change to the likelihood. Note that this cleaning also reduces the uncertainty on  $r$  by a similar order of magnitude. With lower values of  $A_{\text{lens}}$ , corresponding to delensed maps,  $r$  is still consistent with the true value. However, more iterations are required to achieve convergence after delensing, as the reconstruction noise floor is lowered. Delensing does not significantly improve the level of cleaning that it is possible to achieve with the more realistic, non-ideal case. A joint study of delensing and systematics cleaning could become more relevant in the case of higher sensitivity experiments such as PICO [41].

## 8 Conclusions

Systematic effects originating in the instrument pose a major challenge for upcoming CMB experiments seeking to measure the primordial CMB  $B$ -mode of polarization. Many of the

existing techniques for mitigating these effects rely on complex instrument modeling and detailed knowledge of the instrument design. We have presented a detailed case study, implementing a QE approach to carry out cleaning of the CMB  $B$ -mode without detailed prior knowledge of the instrument. We have shown that this QE technique can successfully remove a T to P leakage sourced by a differential detector gain systematic, resulting in a near optimum recovery of the primordial  $B$ -mode and the reduction of the bias on the tensor-to-scalar ratio by  $\sim 2$  orders of magnitude. Our robust implementation builds on the previous work by carrying out systematic recovery and map correction on a full TOD simulation including the effects of a realistic satellite scan strategy, and by the use of newly-derived efficient full-sky estimators. In our recovery and map correction we use a novel Gaussian filter which we find to be an effective alternative in cases where the Wiener filter caused the map correction to fail.

Our case study involved two scenarios. The first scenario, with no noise, beam or lensing, was used to illustrate the absolute limit to the cleaning process in an ideal world when there are no complications. The second scenario provides a more realistic, non-ideal example of the cleaning for a contemporary CMB experiment by using realistic levels of noise and beam comparable to those expected for the *LiteBIRD* instrument [12]. It was necessary to test the iterative cleaning scheme used in our map correction. We carried out this testing using a semi-analytical forecast for the ideal and realistic cleaning. Our cleaning was successful as it was found to be consistent with the forecast

We used our case study to test the conventional wisdom applied in previous studies of this approach [7] and from CMB weak lensing research. We find that in specific cases this wisdom does not hold. For example, previous studies suggest that using EB will result in the best reconstruction of T to P leakage. However, we find that the TB correlations provide the best reconstruction. In some cases using the Wiener filter, the optimum filter that is used in delensing, resulted in divergence when cleaning. The aforementioned Gaussian filter was found to avoid this divergence. These examples, where the conventional wisdom does not apply in the case of systematics cleaning, show the importance of carrying out this case study.

A number of complications exist which will need addressing for this method to be viable that we leave to future work to consider. These complications include the inclusion of foregrounds, and the presence of multiple different systematics. It may be possible to reconstruct and remove these systematics simultaneously. Moreover, it may be possible to break the degeneracy between some cosmological signals and systematics using a QE approach in combination with prior knowledge of the scan strategy.

Despite the additional complications that need to be considered, this detailed study of the QE reconstruction and the improvements made to the iterative cleaning process are an important step towards implementing QEs to reconstruct and remove systematic effects from upcoming CMB surveys. We suggest that this QE technique should be used to compliment traditional systematic correction techniques to diagnose and remove residual contamination in the data not corrected by other methods.

## Acknowledgments

JW and NM are supported by Science and Technology Facilities Council (STFC) studentships. DBT acknowledges support from STFC grant ST/T000341/1 and ST/P000649/1. AR was supported by the ERC Consolidator Grant *CMBSPEC* (No. 725456) as part of the European

Union's Horizon 2020 research and innovation program. MLB acknowledges support from STFC grant ST/T007222/1.

## A Geometric identity

The QEs, as derived in section 3, rely on geometric couplings between the modes of spin-2 and spin-0 fields. Here, we present details of the important geometric terms used in the QE derivations, and derive the identity found in equation (3.9). We begin with the integral of the spherical harmonic terms. In general this term can be written in terms of the Wigner-3j symbols,

$$\int d\hat{\mathbf{n}} {}_{s_1}Y_{l_1 m_1}(\hat{\mathbf{n}}) {}_{s_2}Y_{l_2 m_2}(\hat{\mathbf{n}}) {}_{s_3}Y_{l_3 m_3}(\hat{\mathbf{n}}) = \sqrt{\frac{(2l_1+1)(2l_2+1)(2l_3+1)}{4\pi}} \times \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \begin{pmatrix} l_1 & l_2 & l_3 \\ -s_1 & -s_2 & -s_3 \end{pmatrix}. \quad (\text{A.1})$$

Using this identity we see that the geometric term,  $\pm I_{Mm_2m_1}^{Ll_2l_1}$ , that was introduced in equation (3.4) is explicitly written as

$$\begin{aligned} \pm I_{Mm_2m_1}^{Ll_2l_1} &= \int d\hat{\mathbf{n}} {}_{\pm 2}Y_{LM}(\hat{\mathbf{n}}) {}_{l_2 m_2}Y_{l_2 m_2}(\hat{\mathbf{n}}) {}_{\pm 2}Y_{l_1 m_1}^*(\hat{\mathbf{n}}) \\ &= (-1)^{m_1} \sqrt{\frac{(2L+1)(2l_2+1)(2l_1+1)}{4\pi}} \begin{pmatrix} L & l_2 & l_1 \\ M & m_2 & -m_1 \end{pmatrix} \begin{pmatrix} L & l_2 & l_1 \\ \mp 2 & 0 & \pm 2 \end{pmatrix}. \end{aligned} \quad (\text{A.2})$$

We can use the coupling parity,  $\ell \equiv L + l_2 + l_1$ , and the  $H_{l_2 l_1}^L$  term,

$$H_{l_2 l_1}^L \equiv \sqrt{\frac{(2L+1)(2l_2+1)(2l_1+1)}{4\pi}} \begin{pmatrix} L & l_2 & l_1 \\ -2 & 0 & 2 \end{pmatrix}, \quad (\text{A.3})$$

to simplify equation (A.2) giving

$$\pm I_{Mm_2m_1}^{Ll_2l_1} = (-1)^{m_1} (\pm 1)^\ell H_{l_2 l_1}^L \begin{pmatrix} L & l_2 & l_1 \\ M & m_2 & -m_1 \end{pmatrix}. \quad (\text{A.4})$$

The identity in equation (3.9) can then be derived starting with

$$\sum_{m_2 m_1} \pm I_{Mm_2m_1}^{Ll_2l_1} \pm I_{M'm_2m_1}^{L'l_2l_1} = \sum_{m_2 m_1} (\pm 1)^{\ell+\ell'} H_{l_2 l_1}^L H_{l_2 l_1}^{L'} \begin{pmatrix} L & l_2 & l_1 \\ M & m_2 & -m_1 \end{pmatrix} \begin{pmatrix} L' & l_2 & l_1 \\ M' & m_2 & -m_1 \end{pmatrix}. \quad (\text{A.5})$$

Carrying out the sum on the right hand side and applying the property of the Wigner-3j symbols,

$$\sum_{m_1 m_2} (2L+1) \begin{pmatrix} l_1 & l_2 & L \\ m_1 & m_2 & M \end{pmatrix} \begin{pmatrix} l_1 & l_2 & L' \\ m_1 & m_2 & M' \end{pmatrix} = \delta_{LL'} \delta_{MM'}, \quad (\text{A.6})$$

simplifies the right hand side of (A.5), giving the identity in (3.9),

$$\sum_{m_2 m_1} \pm I_{Mm_2m_1}^{Ll_2l_1} \pm I_{M'm_2m_1}^{L'l_2l_1} = \frac{(H_{l_2 l_1}^L)^2}{2L+1} \delta_{LL'} \delta_{MM'}. \quad (\text{A.7})$$

## References

- [1] C.L. Bennett, A. Banday, K.M. Gorski, G. Hinshaw, P. Jackson, P. Keegstra et al., *Four year COBE DMR cosmic microwave background observations: Maps and basic results*, *Astrophys. J. Lett.* **464** (1996) L1 [[astro-ph/9601067](#)] [[INSPIRE](#)].
- [2] WMAP collaboration, *Nine-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Final Maps and Results*, *Astrophys. J. Suppl.* **208** (2013) 20 [[arXiv:1212.5225](#)] [[INSPIRE](#)].
- [3] PLANCK collaboration, *Planck 2018 results. I. Overview and the cosmological legacy of Planck*, *Astron. Astrophys.* **641** (2020) A1 [[arXiv:1807.06205](#)] [[INSPIRE](#)].
- [4] SIMONS OBSERVATORY collaboration, *The Simons Observatory: Science goals and forecasts*, *JCAP* **02** (2019) 056 [[arXiv:1808.07445](#)] [[INSPIRE](#)].
- [5] M. Remazeilles, A. Rotti and J. Chluba, *Peeling off foregrounds with the constrained moment ILC method to unveil primordial CMB B-modes*, *Mon. Not. Roy. Astron. Soc.* **503** (2021) 2478 [[arXiv:2006.08628](#)] [[INSPIRE](#)].
- [6] B.D. Sherwin and M. Schmittfull, *Delensing the CMB with the Cosmic Infrared Background*, *Phys. Rev. D* **92** (2015) 043005 [[arXiv:1502.05356](#)] [[INSPIRE](#)].
- [7] A.P.S. Yadav, M. Su and M. Zaldarriaga, *Primordial B-mode Diagnostics and Self Calibrating the CMB Polarization*, *Phys. Rev. D* **81** (2010) 063512 [[arXiv:0912.3532](#)] [[INSPIRE](#)].
- [8] PLANCK collaboration, *Planck 2018 results. VIII. Gravitational lensing*, *Astron. Astrophys.* **641** (2020) A8 [[arXiv:1807.06210](#)] [[INSPIRE](#)].
- [9] V. Gluscevic, D. Hanson, M. Kamionkowski and C.M. Hirata, *First CMB Constraints on Direction-Dependent Cosmological Birefringence from WMAP-7*, *Phys. Rev. D* **86** (2012) 103529 [[arXiv:1206.5546](#)] [[INSPIRE](#)].
- [10] POLARBEAR collaboration, *POLARBEAR Constraints on Cosmic Birefringence and Primordial Magnetic Fields*, *Phys. Rev. D* **92** (2015) 123509 [[arXiv:1509.02461](#)] [[INSPIRE](#)].
- [11] BICEP2 and KECK ARRAY collaborations, *BICEP2/Keck Array IX: New bounds on anisotropies of CMB polarization rotation and implications for axionlike particles and primordial magnetic fields*, *Phys. Rev. D* **96** (2017) 102003 [[arXiv:1705.02523](#)] [[INSPIRE](#)].
- [12] M. Hazumi et al., *LiteBIRD: A Satellite for the Studies of B-Mode Polarization and Inflation from Cosmic Background Radiation Detection*, *J. Low Temp. Phys.* **194** (2019) 443 [[INSPIRE](#)].
- [13] W. Hu, M.M. Hedman and M. Zaldarriaga, *Benchmark parameters for CMB polarization experiments*, *Phys. Rev. D* **67** (2003) 043004 [[astro-ph/0210096](#)] [[INSPIRE](#)].
- [14] D. O’Dea, A. Challinor and B.R. Johnson, *Systematic errors in cosmic microwave background polarization measurements*, *Mon. Not. Roy. Astron. Soc.* **376** (2007) 1767 [[astro-ph/0610361](#)] [[INSPIRE](#)].
- [15] BICEP2 collaboration, *BICEP2 III: Instrumental Systematics*, *Astrophys. J.* **814** (2015) 110 [[arXiv:1502.00608](#)] [[INSPIRE](#)].
- [16] POLARBEAR collaboration, *A Measurement of the Cosmic Microwave Background B-Mode Polarization Power Spectrum at Sub-Degree Scales with POLARBEAR*, *Astrophys. J.* **794** (2014) 171 [[arXiv:1403.2369](#)] [[INSPIRE](#)].
- [17] J. Delabrouille and J.-F. Cardoso, *Diffuse source separation in cmb observations*, in *Data Analysis in Cosmology*, V. J. Martinez, E. Saar, E. M. Gonzales and M. J. Pons-Borderia, eds., Springer Berlin Heidelberg, Berlin, Heidelberg, Germany (2009), pp. 159–205.
- [18] K. Ichiki, *CMB foreground: A concise review*, *PTEP* **2014** (2014) 06B109 [[INSPIRE](#)].

- [19] N. McCallum, D.B. Thomas, M.L. Brown and N. Tessore, *Spin characterization of systematics in CMB surveys – a comprehensive formalism*, *Mon. Not. Roy. Astron. Soc.* **501** (2021) 802 [[arXiv:2008.00011](#)] [[INSPIRE](#)].
- [20] J. Carron, A. Lewis and A. Challinor, *Internal delensing of Planck CMB temperature and polarization*, *JCAP* **05** (2017) 035 [[arXiv:1701.01712](#)] [[INSPIRE](#)].
- [21] A. Lewis and A. Challinor, *Weak gravitational lensing of the CMB*, *Phys. Rept.* **429** (2006) 1 [[astro-ph/0601594](#)] [[INSPIRE](#)].
- [22] PLANCK collaboration, *Planck 2015 results. XVI. Isotropy and statistics of the CMB*, *Astron. Astrophys.* **594** (2016) A16 [[arXiv:1506.07135](#)] [[INSPIRE](#)].
- [23] T. Namikawa et al., *Atacama Cosmology Telescope: Constraints on cosmic birefringence*, *Phys. Rev. D* **101** (2020) 083527 [[arXiv:2001.10465](#)] [[INSPIRE](#)].
- [24] BICEP2 and KECK ARRARY collaborations, *BICEP2/Keck Array IX: New bounds on anisotropies of CMB polarization rotation and implications for axionlike particles and primordial magnetic fields*, *Phys. Rev. D* **96** (2017) 102003 [[arXiv:1705.02523](#)] [[INSPIRE](#)].
- [25] J. Williams, A. Rotti and R. Battye, *Constraining cosmic polarization rotation and implications for primordial B-modes*, *JCAP* **09** (2020) 006 [[arXiv:2006.04899](#)] [[INSPIRE](#)].
- [26] W. Hu, *Weak lensing of the CMB: A harmonic approach*, *Phys. Rev. D* **62** (2000) 043007 [[astro-ph/0001303](#)] [[INSPIRE](#)].
- [27] C.G.R. Wallis, M.L. Brown, R.A. Battye and J. Delabrouille, *Optimal scan strategies for future CMB satellite experiments*, *Mon. Not. Roy. Astron. Soc.* **466** (2017) 425 [[arXiv:1604.02290](#)] [[INSPIRE](#)].
- [28] EPIC collaboration, *Study of the Experimental Probe of Inflationary Cosmology (EPIC)-Intermediate Mission for NASA’s Einstein Inflation Probe*, [arXiv:0906.1188](#) [[INSPIRE](#)].
- [29] BICEP2 collaboration, *BICEP2 III: Instrumental Systematics*, *Astrophys. J.* **814** (2015) 110 [[arXiv:1502.00608](#)] [[INSPIRE](#)].
- [30] N.J. Miller, M. Shimon and B.G. Keating, *CMB Beam Systematics: Impact on Lensing Parameter Estimation*, *Phys. Rev. D* **79** (2009) 063008 [[arXiv:0806.3096](#)] [[INSPIRE](#)].
- [31] ABS collaboration, *Modulation of cosmic microwave background polarization with a warm rapidly rotating half-wave plate on the Atacama B-Mode Search instrument*, *Rev. Sci. Instrum.* **85** (2014) 024501 [[arXiv:1310.3711](#)] [[INSPIRE](#)].
- [32] A. Kusaka et al., *Results from the Atacama B-mode Search (ABS) Experiment*, *JCAP* **09** (2018) 005 [[arXiv:1801.01218](#)] [[INSPIRE](#)].
- [33] J. Bock, A. Cooray, S. Hanany, B. Keating, A. Lee, T. Matsumura et al., *The Experimental Probe of Inflationary Cosmology (EPIC): A Mission Concept Study for NASA’s Einstein Inflation Probe*, [arXiv:0805.4207](#) [[INSPIRE](#)].
- [34] D.B. Thomas, N. McCallum and M.L. Brown, *Sky rotation in ground-based cosmic microwave background experiments*, [arXiv:2102.02284](#) [[INSPIRE](#)].
- [35] K.M. Gorski, E. Hivon, A.J. Banday, B.D. Wandelt, F.K. Hansen, M. Reinecke et al., *HEALPix — A Framework for high resolution discretization, and fast analysis of data distributed on the sphere*, *Astrophys. J.* **622** (2005) 759 [[astro-ph/0409513](#)] [[INSPIRE](#)].
- [36] A. Lewis, A. Challinor and A. Lasenby, *Efficient computation of CMB anisotropies in closed FRW models*, *Astrophys. J.* **538** (2000) 473 [[astro-ph/9911177](#)] [[INSPIRE](#)].
- [37] M. Hazumi et al., *LiteBIRD: A Satellite for the Studies of B-Mode Polarization and Inflation from Cosmic Background Radiation Detection*, *J. Low Temp. Phys.* **194** (2019) 443 [[INSPIRE](#)].

- [38] D.B. Thomas, N. McCallum and M.L. Brown, *Controlling systematics in ground-based CMB surveys with partial boresight rotation*, *Mon. Not. Roy. Astron. Soc.* **491** (2020) 1960 [[arXiv:1905.12647](#)] [[INSPIRE](#)].
- [39] S. Hamimeche and A. Lewis, *Likelihood Analysis of CMB Temperature and Polarization Power Spectra*, *Phys. Rev. D* **77** (2008) 103013 [[arXiv:0801.0554](#)] [[INSPIRE](#)].
- [40] N. Katayama and E. Komatsu, *Simple foreground cleaning algorithm for detecting primordial B-mode polarization of the cosmic microwave background*, *Astrophys. J.* **737** (2011) 78 [[arXiv:1101.5210](#)] [[INSPIRE](#)].
- [41] M. Alvarez et al., *PICO: Probe of Inflation and Cosmic Origins*, [arXiv:1908.07495](#) [[INSPIRE](#)].